

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2004, 09:12:48 ; Search time 34 Seconds

(without alignments)
1607.317 Million cell updates/sec

Title: US-09-990-415A-2

Perfect score: 1345
Sequence: 1 MESKMGELPLDINIOEPRWD.....PANAIPLMDTEKDFLK 261

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

```
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1345	100.0	261	10 US-09-990-415A-2	Sequence 2, App1
2	1345	100.0	265	9 US-09-867-550-1344	Sequence 1344, Ap
3	1089	81.0	322	10 US-09-990-415A-8	Sequence 8, App1
4	1089	81.0	322	15 US-10-014-338-2	Sequence 2, App1
5	876	65.1	261	15 US-10-014-338-4	Sequence 4, App1
6	753	56.0	322	10 US-09-990-415A-6	Sequence 2, App1
7	436	32.4	186	12 US-10-094-749-2165	Sequence 2165, Ap
8	368.5	27.4	77	10 US-09-990-415A-4	Sequence 4, App1
9	312	23.2	127	12 US-10-264-237-2468	Sequence 2468, Ap
10	143.5	10.7	337	12 US-10-415-378-1	Sequence 1, App1
11	143	10.6	54	15 US-10-106-698-6940	Sequence 6940, Ap
12	96	7.1	836	11 US-09-972-708-9	Sequence 9, App1
13	91	6.8	496	12 US-10-104-047-3006	Sequence 3006, Ap
14	89.5	6.7	344	15 US-10-306-133-3	Sequence 3, App1
15	88.5	6.6	344	10 US-09-978-295A-523	Sequence 523, App

	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
	88.5	88.5	88.5	88.5	88.5	88.5	88.5	88.5	88.5	88.5	88.5	88.5	88.5	88.5	88.5	88.5	88.5	88.5	88.5	88.5	88.5	88.5	88.5	88.5	88.5	88.5	88.5	88.5	88.5	88.5
	6.6	6.6	6.6	6.6	6.6	6.6	6.6	6.6	6.6	6.6	6.6	6.6	6.6	6.6	6.6	6.6	6.6	6.6	6.6	6.6	6.6	6.6	6.6	6.6	6.6	6.6	6.6	6.6	6.6	6.6
	344	344	344	344	344	344	344	344	344	344	344	344	344	344	344	344	344	344	344	344	344	344	344	344	344	344	344	344	344	344
	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	US-09-966-546-4	US-09-966-546-6	US-09-978-697-523	US-09-966-545-6	US-09-966-545-6	US-09-966-545-6	US-09-966-545-6	US-09-966-545-6	US-09-966-545-6	US-09-966-545-6	US-09-966-545-6	US-09-966-545-6	US-09-966-545-6	US-09-966-545-6	US-09-966-545-6	US-09-966-545-6	US-09-966-545-6	US-09-966-545-6	US-09-966-545-6	US-09-966-545-6	US-09-966-545-6	US-09-966-545-6	US-09-966-545-6	US-09-966-545-6	US-09-966-545-6	US-09-966-545-6	US-09-966-545-6	US-09-966-545-6	US-09-966-545-6	US-09-966-545-6
	Sequence 4, App1	Sequence 6, App1	Sequence 523, App	Sequence 4, App1	Sequence 6, App1	Sequence 523, App	Sequence 4, App1	Sequence 6, App1	Sequence 523, App	Sequence 4, App1	Sequence 6, App1	Sequence 523, App	Sequence 4, App1	Sequence 6, App1	Sequence 523, App	Sequence 4, App1	Sequence 6, App1	Sequence 523, App	Sequence 4, App1	Sequence 6, App1	Sequence 523, App	Sequence 4, App1	Sequence 6, App1	Sequence 523, App	Sequence 4, App1	Sequence 6, App1	Sequence 523, App	Sequence 4, App1	Sequence 6, App1	Sequence 523, App

ALIGNMENTS

```
RESULT 1
US-09-990-415A-2
; Sequence 2, Application US/09990415A
; Patent No. US20020165182A1
; GENERAL INFORMATION:
; APPLICANT: Pharmacia AB
; TITLE OF INVENTION: Protein Cluster 1
; FILE REFERENCE: 00349
; CURRENT APPLICATION NUMBER: US/09/990,415A
; CURRENT FILING DATE: 2001-11-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 261
; TYPE: PRT
; ORGANISM: human
US-09-990-415A-2
```

Query Match 100.0%; Score 1345; DB 10; Length 261;
Best Local Similarity 100.0%; Pred. No. 2.2e-141;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MESKMGELPLDINIOEPRWDSTFLGRARHFTVTDPRNLLSGAOLEASRNIVQNYRAG	60
DB	1	MESKMGELPLDINIOEPRWDSTFLGRARHFTVTDPRNLLSGAOLEASRNIVQNYRAG	60
QY	61	VTPRGITEDOLMRAKYVYDSAFHPDTEKVVLLIGRNSAQVPMNNTTIGCMLTFPRKPTV	120
DB	61	VTPRGITEDOLMRAKYVYDSAFHPDTEKVVLLIGRNSAQVPMNNTTIGCMLTFPRKPTV	120
QY	121	VFMQWVQSFNAIVYNSRSGDPTITVROJGTAVASATGAVAATAGLAKSLTKHLPPLV	180
DB	121	VFMQWVQSFNAIVYNSRSGDPTITVROJGTAVASATGAVAATAGLAKSLTKHLPPLV	180
QY	181	RFVFPFAVAANACINIPLMKQREIQVGI PVADERQGLGYSVTNAKQIGIFQVVISRICMA	240
DB	181	RFVFPFAVAANACINIPLMKQREIQVGI PVADERQGLGYSVTNAKQIGIFQVVISRICMA	240

Qy	241	IPAMAI PPLIMDTLEKKDFLK	261
Db	241	IPAMAI PPLIMDTLEKKDFLK	261

RESULT 2

US-09-867-550-1344
Sequence 1344, Application US/09867550
Patent No. US2002008206A1
GENERAL INFORMATION:
APPLICANT: Leach, Martin D.
APPLICANT: Mehraban, Foad,
APPLICANT: Conley, Pamela
APPLICANT: Law, Debbie
TITLE OF INVENTION: No. US2002008206A1 Polynucleotides from Atherogenic Cells and
FILE REFERENCE: 21402-013 (Cura-113)
CURRENT APPLICATION NUMBER: US/09/867,550
CURRENT FILING DATE: 2001-09-20
PRIOR APPLICATION NUMBER: USSN 60/208,427
PRIOR FILING DATE: 2000-05-30
NUMBER OF SEQ ID NOS: 2125
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1344
LENGTH: 266
TYPE: PRT
ORGANISM: Homo sapiens
US-09-867-550-1344

Query Match	100.0%	Score 1345:	DB 9:	Length 266:
Best Local Similarity	100.0%	Pred. No. 2.2e-141:		
Matches 261:	Conservative 0:	Mismatches 0:	Indels 0:	Gaps 0

Qy	1	MEKMELEPLDINIOBPBNDOSTFLGRABHFTYDPPENLLISGOLBASINIVQNRAG	60
Db	1	MEKMELEPLDINIOBPBNDOSTFLGRABHFTYDPPENLLISGOLBASINIVQNRAG	60
Qy	61	VVTGGITBDOLWRAKRVYDSAFHPDTEKVVYIGMSAQVPNNMTITGCMLEFYRKPTV	120
Db	61	VVTGGITBDOLWRAKRVYDSAFHPDTEKVVYIGMSAQVPNNMTITGCMLEFYRKPTV	120
Qy	121	VPMQWNOGSNNAIVNYSNRSGDPTPIVROLGPAVYSATTTGAVATLGLKLTTHLPLVNG	180
Db	121	VPMQWNOGSNNAIVNYSNRSGDPTPIVROLGPAVYSATTTGAVATLGLKLTTHLPLVNG	180
Qy	181	RFVFPAAVAAANCINIPLMRORELQYGIIVADAGORLIGYSTAKOGIPQVVISRICMA	240
Db	181	RFVFPAAVAAANCINIPLMRORELQYGIIVADAGORLIGYSTAKOGIPQVVISRICMA	240
Qy	241	IPMAAIPLIMDLEKKDFLK	261
Db	241	IPMAAIPLIMDLEKKDFLK	261

RESULT 3

```

US-09-990-415A-8
Sequence 8, Application US/0990415A
Patent No. US20020165182A1
GENERAL INFORMATION:
APPLICANT: Pharmacia AB
TITLE OF INVENTION: Protein Cluster I
FILE REFERENCE: 00349
CURRENT APPLICATION NUMBER: US/09/990,415A
CURRENT FILING DATE: 2001-11-21
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn version 3.0
SEQ ID NO 8
LENGTH: 322
TYPE: PRT
ORGANISM: human
FEATURE:
NAME/KEY: misc feature

```

! LOCATION: (25)..(25)
; OTHER INFORMATION: Xaa=A,T,G or C
US-09-990-415A-8

Query Match	81.0%;	Score 1089;	DB 10;	Length 322;
Best Local Similarity	78.5%;	Pred. No. 1.1e-112;		
Matches 201; Conservative	28;	Mismatches 27;	Indels 0;	Gaps 0

Qy		6	GELPLDINIOERPMOSFLEGRARHFVTPTPRNLLISGAOLEASRNIVONRYAGVTPG	65
Dd		3	GELPPIINIKERMOOSTIFIGRAHNFVTPTPRNILLTNEDEBARKIIYHDYRGCIYPBG	62
Qy		66	ITEDOLMRKAVYYDSAFPHDTGEKKVLLIRMSAOPNNMTITGCMLTFYRKTPTVFWM	1250
Dd		63	LTENELMRKAYIYDBAFPHDTGEKKMILLGRMSAOVPNMNTITGCMMTFYRTTPAVLFWOM	1220
Qy		126	VNOSEPAIYNVSNSGGDPITVROLGFAYSATTAATGAVALTGLSLTKHLPELVGRFVFP	1890
Dd		123	INOSFAAVNNTNRSGDALPVINEGLGVARSATTGAVALTGMAALTKHSFPLGRVFP	1880
Qy		186	AAVAANNCINIEPLMEORELOVGI PVADBAQRGLGYSTVAAKQGI FOYVISRICMAIPAMA	2450
Dd		183	AAVAANNCINIEPLMEORELKVGIPVTDENGNRLGESANAQAQTQQVVVSRIIMAAPGMA	2420
Qy		246	IPELMIDTLKEKDPLX 261	
Dd		243	IPELMIDTLKEKAFLX 258	

RESULT 4

```

US-10-014-338-2
Sequence 2, Application US/10014338
Publication No. US20030092614A1
GENERAL INFORMATION:
APPLICANT: Heatch, et al.
TITLE OF INVENTION: ADPI-41, A NOVEL PROTEIN ISOLATED FROM BRAIN TISSUE HOMOGENATE A
TITLE OF INVENTION: USES THEREFOR
FILE REFERENCE: 9195-077
CURRENT APPLICATION NUMBER: US//10/014,338
CURRENT FILING DATE: 2002-05-01
PRIOR APPLICATION NUMBER: 10/014,338
PRIOR FILING DATE: 2001-12-10
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 322
TYPE: PRT
ORGANISM: Homo sapiens
US-10-014-338-2

```

Query Match	81.0%	Score 1089;	DB 15;	Length 322;
Best Local Similarity	78.5%	Pred. No. 1,1e-112;		
Matches 201; Conservative	28;	Mismatches 27;	Indels 0;	Gaps 0

[illegible]

RESULT 5
US-10-014-338-4
; Sequence 4, Application US/10014338
; Publication No. US20030092614A1
; GENERAL INFORMATION:
; APPLICANT: Herath, et al.
; TITLE OF INVENTION: ADP1-41, A NOVEL PROTEIN ISOLATED FROM BRAIN TISSUE HOMOGENATE AN
; TITLE OF INVENTION: USES THEREFOR
; FILE REFERENCE: 9195-077
; CURRENT APPLICATION NUMBER: US/10/014,338
; CURRENT FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: 10/014,338
; PRIOR FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-014-338-4

Query Match 65.1%; Score 876; DB 15; Length 261;
Best Local Similarity 79.6%; Pred. No. 4,5e-89;
Matches 156; Conservative 25; Mismatches 15; Indels 0; Gaps 0;

QY 6 GELPLDINIQEPKWDSTFLGRARHFFVTVDPRNLLLSGAQLBASNIYONTRAGVTTG 65
DB 3 GELPPINIKERPMQSTFLGRANHFVTVDPRNLLLTNEQLSARKIVHDYRQGIIVPG 62
QY 66 ITEDQIMRAKYVYDFAFHDPDTEKXVLIGRMSAQVPMNMTITGCMLEFRKPTVVFQW 125
DB 63 LLENELMRKAYVYDFAFHDPDTEKMLIGRMSAQVPMNMTITGCMLEFRTPAVLFWQ 122
QY 126 VNOSEFNAIVYNSRSGDPTITVROLGTAVYSATGAVATAGLTKLPLVGRFVVF 185
DB 123 INQSFNAVYNTNRSDDALITVNELGTAIVSYTTGAVATAGLTKVSVLGRFVVF 182
QY 186 AAVAAANCINIFLMRQ 201
DB 183 AAVAAANCINIFLMRQ 198

RESULT 6
US-09-990-415A-6
; Sequence 6, Application US/09990415A
; Patent No. US20020165182A1
; GENERAL INFORMATION:
; APPLICANT: Pharmacia AB
; TITLE OF INVENTION: Proteoln Cluster I
; FILE REFERENCE: 00349
; CURRENT APPLICATION NUMBER: US/09/990,415A
; CURRENT FILING DATE: 2001-11-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 322
; TYPE: PRT
; ORGANISM: human
US-09-990-415A-6

Query Match 56.0%; Score 753; DB 10; Length 322;
Best Local Similarity 59.0%; Pred. No. 3,2e-75;
Matches 147; Conservative 37; Mismatches 65; Indels 0; Gaps 0;

QY 13 NIOEPKWDSTFLGRARHFFVTVDPRNLLLSGAQLBASNIYONTRAGVTTGITEQW 72
DB 9 NIDAPKWDSTFLGRARHFFVTVDPRNLLLSGAQLBASNIYONTRAGVTTGITEQW 68
QY 73 RAKYVYDFAFHDPDTEKXVLIGRMSAQVPMNMTITGCMLEFRKPTVVFQWVNSFNA 132
DB 69 YAKKYDFAFHDPDTEKXVLIGRMSAQVPMNMTITGCMLEFRKPTVVFQWVNSFNA 128

QY 133 IYVNSRSGDPTITVROLGTAVYSATGAVATAGLTKLPLVGRFVFPAAVAAAN 192
DB 129 LVNVTNRNAASTVSQMLSTFTATTAVATAGVGNMILTKKAPLVGRWVFPAAVAAAN 188
QY 193 CINIPKRELEQVQIPVADBGRLGYSVTAAKQIFQVVISRICALPAAAIPLIND 252
DB 189 CNIPKRELEQVQIPVADBGRLGYSVTAAKQIFQVVISRICALPAAAIPLIND 248
QY 253 TLEKDFLK 261
DB 249 RLEKLFHQ 257

RESULT 7
US-10-094-749-2165
; Sequence 2165, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHITO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOKYU
; APPLICANT: NAGAHARA, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 08435/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2165
; LENGTH: 186
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2165

Query Match 32.4%; Score 436; DB 12; Length 186;
Best Local Similarity 56.5%; Pred. No. 3,3e-40;
Matches 83; Conservative 22; Mismatches 42; Indels 0; Gaps 0;

QY 3 SNKGELPLDINIQEPKWDSTFLGRARHFFVTVDPRNLLLSGAQLBASNIYONTRAGV 62
DB 17 SKGEADLSGFNIDAPKWDSTFLGRARHFFVTVDPRNLLLSGAQLBASNIYONTRAGV 76
QY 63 TPGITEDQIMRAKYVYDFAFHDPDTEKXVLIGRMSAQVPMNMTITGCMLEFRKPTVVF 122
DB 77 PPGTVEQLLYAKKYDFAFHDPDTEKXVLIGRMSAQVPMNMTITGCMLEFRKPTVVF 136
QY 123 WQVNSFNAIVYNSRSGDPTITVRO 149
DB 137 WQVNSFNAIVYNTNRNAASTVS 163

RESULT 8
US-09-990-415A-4
; Sequence 4, Application US/09990415A

```

; Patent No. US20020165182A1
; GENERAL INFORMATION:
; APPLICANT: Pharmacia AB
; TITLE OF INVENTION: Protein Cluster I
; FILE REFERENCE: 00349
; CURRENT APPLICATION NUMBER: US/09/990,415A
; CURRENT FILING DATE: 2001-11-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 77
; TYPE: PRT
; ORGANISM: human
US-09-990-415A-4

Query Match      27.4%; Score 368.5; DB 10; Length 77;
Best Local Similarity 57.5%; Pred. No. 3,1e-33;
Matches 77; Conservative 0; Mismatches 0; Indels 57; Gaps 1;

Qy      1 MESMGELPLDINTQEPKMDSTFLGRARHFFVTYDPRNLLSGAQLASRNIVQNYRAG 60
Db      1 MESMGELPLDINTQEPKMDSTFLGRARHFFVTYDPRNLLSGAQLASRNIVQNYRAG 56
Qy      61 VVTGRIEDQMRKAYVYDSAFHPTGKXVVLGRMSAQVPMNTTIGCMILTFKRTPTV 120
Db      57 -----YRKTPTV 63

Qy      121 VFMQVNSFNAIV 134
Db      64 VFMQVNSFNAIV 77

RESULT 9
US-10-264-237-2468
; Sequence 2468, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA313P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 2468
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-237-2468

Query Match      23.2%; Score 312; DB 12; Length 127;
Best Local Similarity 100.0%; Pred. No. 1.3e-26;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      199 MRORELOVIGPVADAGORLGYSYTAAGKGFQVVISRICMAIPMAIPLIMTLEKCD 258
Db      1 MRORELOVIGPVADAGORLGYSYTAAGKGFQVVISRICMAIPMAIPLIMTLEKCD 60
Qy      259 FLK 261
Db      61 FLK 63

RESULT 10
US-10-415-378-1
; Sequence 1, Application US/10415378
; Publication No. US20040014945A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; TANG, Y. TOM

```

```

; APPLICANT: YUE, Henry; NGUYEN, Daniel B.;
; APPLICANT: HAFALIA, April J.A.; ELLIOTT, Vicki S.;
; APPLICANT: LU, Yan; CHAMLA, Narinder K.;
; APPLICANT: YAO, Monique G.; BAUGHN, Mariah R.;
; APPLICANT: GANDHI, Ameena R.; DING, Li;
; APPLICANT: SANJANMALA, Madhusudan M.; RAMKUMAR, Jayalaxmi;
; APPLICANT: ARVIZU, Chandra S.; GIEZEN, Kimberly J.;
; APPLICANT: LAL, Preeti G.; AZIMZAI, Yalda;
; APPLICANT: KHAN, Farrah A.; THANGAVELU, Kavitha;
; APPLICANT: THORNTON, Michael B.; LU, Dyung Alma M.;
; APPLICANT: TRIBOUTLEY, Catherine M.; WARREN, Bridget A.;
; APPLICANT: ISON, H. Craig; DAS, Debopriya;
; APPLICANT: RAUMANN, Brigitte E.; POLICKI, Jennifer L.;
; APPLICANT: KEARNEY, Liam
; TITLE OF INVENTION: TRANSPORTERS AND ION CHANNELS
; FILE REFERENCE: PI-0270 USN
; CURRENT APPLICATION NUMBER: US/10/415,378
; CURRENT FILING DATE: 2003-05-07
; PRIOR APPLICATION NUMBER: PCT/US01/46055
; PRIOR FILING DATE: 2001-10-27
; PRIOR APPLICATION NUMBER: US 60/250,790
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 60/252,232
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/249,661
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: US 60/247,673
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/245,904
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: US 60/243,989
; PRIOR FILING DATE: 2000-10-27
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040014945A1 1626101CD1
US-10-415-378-1

Query Match      10.7%; Score 143.5; DB 12; Length 337;
Best Local Similarity 22.1%; Pred. No. 3.4e-07;
Matches 57; Conservative 56; Mismatches 128; Indels 17; Gaps 6;

Qy      16 EPR---W--DQSTFLGRARHFFVTYDPRNLLSGAQLASRNIVQNYRAGVTPGITEDQ 70
Db      26 EPNVRFWITERQSFIRFLQWTELDPTNVFISVESIENSRLCT--NEDVSSPASADQR 84
Qy      71 LMRKAYVYDSAFHPTGKXVVLGRMSAQVPMNTTIGCMILTFYKRTPTVFMQVNSF 130
Db      85 IQEAKRSGLATVHPDSSNLLPKLFRPAFLPFAAPTYFLSMTPLKGIKSVILPQVFLCAY 144
Qy      131 NAIYVNSNRSGDPTTYRQLGTAVSATTGAVATAGLKLTKLP-----PLVGRFV 183
Db      145 MAAFNSIN--GNRSYTCPLERSLMA--GAVASSTFLGVTPQVQKXGLTGFWIRRL 200
Qy      184 PFAVAANCINIPMRQRELOVIGPVADAGORLGYSYTAAGKGFQVVISRICMAIPA 243
Db      201 PVIFLVQASGMNVMSRSLISIKGIAYVMDKGVNLSHISIACTAVAKETLASRIVLFGTS 260
Qy      244 MAIPLIMTLEKCDFLK 261
Db      261 ALIPEVFTYFFKRTQYFR 278

RESULT 11
US-10-106-698-6940
; Sequence 6940, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:

```

APPLICANT: Ruben et al.
TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
FILE REFERENCE: PA005P1
CURRENT APPLICATION NUMBER: US/10/106,698
CURRENT FILING DATE: 2002-03-27
PRIOR APPLICATION NUMBER: PCT/US00/26524
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US 60/157,137
PRIOR FILING DATE: 1999-09-29
PRIOR APPLICATION NUMBER: US 60/163,280
PRIOR FILING DATE: 1999-11-03
NUMBER OF SEQ ID NOS: 8564
SOFTWARE: PatentIn Ver. 3.0
SEQ ID NO: 6940
LENGTH: 54
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (14)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: MISC_FEATURE
LOCATION: (28)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: MISC_FEATURE
LOCATION: (35)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-6940

Query Match
Best Local Similarity 10.6%; Score 143; DB 15; Length 54;
Matches 27; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 15 QEPMDQSTFLGRARHFTVTDPRNLLSG 44
5 QEPMDQSTXPGRARHFTVTDPRNLLSG 34

RESULT 12
US-09-972-708-9
Sequence 9, Application US/09972708
Publication No. US20030059871A1
GENERAL INFORMATION:
APPLICANT: Immunex Corporation
APPLICANT: Cosman, David J.
APPLICANT: Mosley, Bruce A.
APPLICANT: Bird, Timothy A.
APPLICANT: Dubose, Robert F.
APPLICANT: Wiley, Steven R.
TITLE OF INVENTION: HEMATOPOIETIN RECEPTORS HPRI AND HPR2
FILE REFERENCE: 3160-B
CURRENT APPLICATION NUMBER: US/09/972,708
CURRENT FILING DATE: 2001-10-05
NUMBER OF SEQ ID NOS: 29
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 9
LENGTH: 836
TYPE: PRT
ORGANISM: Homo sapiens
US-09-972-708-9

Query Match
Best Local Similarity 7.1%; Score 96; DB 11; Length 836;
Matches 54; Conservative 34; Mismatches 93; Indels 36; Gaps 12;

Db 61 VTPGTEQLMRAK--YVYDSAFHPTGKVL--IGRMSAO--VPMNMTTGMLTF 113
503 IYTP-LYQDTMGSSQHYVAYSQMSAHBELHKLIGTKWQLEWVPEPELGSPLTH 561
114 YRKTPTVFM-QWVNSFNAIVYNSRS---GDTPT--VQQLGTAVYASATGAVATA 165
562 Y-----TIFWTNQNQSFSLINASSRGFVHGLBPASLVYHILMAASQAGATNSTVLTL 616

166 LGLKSLTKHLPELVGRF---VPEAAVAAANCINIPLMRRELQVGIPIVADSGORLGSV 222
617 MFLTPREGSELHITLPLFGILLITLCTGAWLCCSPNKNPIMPSPV--DPAHSSLGWV 674

Db 223 -TAAKQIFQVVISRICMAIPAMAIPPLIMDLERKD 258
675 PTIMEDDAFQ-----LPGLGTPPTIKLTVLEED 702

RESULT 13
US-10-104-047-3006
Sequence 3006, Application US/10104047
Publication No. US20030236392A1
GENERAL INFORMATION:
APPLICANT: HELIX RESEARCH INSTITUTE
TITLE OF INVENTION: NO. US20030236392A1 full length cDNA
FILE REFERENCE: H1-A0105
CURRENT APPLICATION NUMBER: US/10/104,047
CURRENT FILING DATE: 2002-03-25
PRIOR APPLICATION NUMBER:
PRIOR FILING DATE:
NUMBER OF SEQ ID NOS: 4096
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 3006
LENGTH: 496
TYPE: PRT
ORGANISM: Homo sapiens
US-10-104-047-3006

Query Match
Best Local Similarity 6.8%; Score 91; DB 12; Length 496;
Matches 47; Conservative 22; Mismatches 83; Indels 24; Gaps 7;

Db 12 INIOEPMDQSTFLGRARHFTVTDPRNLLSGAQLASRN-----IVQNYRAGVVT 63
67 LGIINPRGGQTYAOKLQGRVTMTSDTATVNMELSLRSDTAIVYCVSAAYCSGDCY 126

Db 64 PGITEDQLM-PAKYVYDASAFHPTGKVLIGRMSAQVPMNMTTGCMILTFRKPTTVF 122
127 PGIFD--LWGRGTLYVSAASP-TSPKVPFLSLCSTQPDGNVIVACLVGQFPQEPISVT 183

Db 123 WQVWNSFNAIVNY---SNRSGDPTITVROQLGTAVYASATGAVATLGLKSLTKHL 175
184 WSESQGVTH-RNFPQSDASGDLTYSQ-----TLPATQCLAKSVCYCHV 230

RESULT 14
US-10-306-133-3
Sequence 3, Application US/10306133
Publication No. US20030100485A1
GENERAL INFORMATION:
APPLICANT: Lel, Preeti
Guegler, Karl J.
TITLE OF INVENTION: HUMAN NEUROTRIMIN HOMOLOG
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSER: Inocyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/306,133
FILING DATE: 27-Nov-02 US20030100485A1-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/009,841

ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
RECEIPT/DOCKET NUMBER: PF-0463 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 344 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 755185
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-306-133-3

Query Match 6.7%; Score 89.5; DB 15; Length 344;
Best Local Similarity 24.8%; Pred.No. 0.37;
Matches 52; Conservative 28; Mismatches 95; Indels 35; Gaps 10;

QY 2 ESKKGEPLDINIEPRM-DOSTFLGRARHFTVTDPRNLLSGAOLEASRNT--VQNYR 58
DB 51 ESATLRCTIDRVRVAMLNSTLYAGNDKMC-DPRVLLSTQYQYSIEIONVVD 109
QY 59 AGVTPGTEQDLRAKVVYDSAFHPDGEKVLIGRMSAV-----PMNNTIT 107
DB 110 EGPTSCVQDN-----HFKT-SRVHLYQVSPTIVSISDISINEENISLT 156
QY 108 GCMLEFRKTPYVFWQWNSFNAIVNSRSDPTIVRQLGTAVYSATGAVATAG 167
DB 157 -CIAT-GRPEPTV-MRHSIPKAVGVSEDEYELIGITRQSGEYCSANDVAAPVR 213
QY 168 LKSLTKLPLV---GRFVFAVAALNC 193
DB 214 RVNTVNPYPYISEAKGTGVPVQKGLTQC 243

RESULT 15
US-09-978-295A-523
Sequence 523, Application US/09978295A
Patent No. US20020156006A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerltsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gutney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kluvin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey

APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630P1C11
CURRENT APPLICATION NUMBER: US/09/978,295A
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049

;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081071
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081195
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081203
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081229
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081955
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081817
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081819
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081952
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081838
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/082568
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082569
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082704
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082804
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082700
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082797
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082796
;; PRIOR FILING DATE: 1998-04-23
;; PRIOR APPLICATION NUMBER: 60/083336
;; PRIOR FILING DATE: 1998-04-27
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/083392
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083495
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083496
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083499
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083545
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083554
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083558
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083559
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083500
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083742
;; PRIOR FILING DATE: 1998-04-30
;; PRIOR APPLICATION NUMBER: 60/084366
;; PRIOR FILING DATE: 1998-05-05
;; PRIOR APPLICATION NUMBER: 60/084414
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084441
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084637
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084639
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084640
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084598
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07

;; PRIOR APPLICATION NUMBER: 60/084627
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084643
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/085339
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085338
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085323
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085582
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085700
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085689
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085580
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085573
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085704
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085697

Query Match 6.6%; Score 88.5; DB 10; Length 344;
Best Local Similarity 24.8%; Pred. No. 0.48;
Matches 52; Conservative 27; Mismatches 96; Indels 35; Gaps 10;

QY 2 ESKMBELPIDINIGPRW-DOSTPLGRARHFTYDPRNLISGQLEASRNI--YQNYR 58
DB 51 ESATLRCTIDNRVTAVAMLRSTLIYAGNDKXCL-DPRVLLSNTQTOYISIQNDVYD 109
QY 59 AGVTPGITEQMLRAKYVDSAFHPDTEKVVLLIGMSAGV-----PMMIT 107
DB 110 BGPYCSVQTDN-----HPKT-SRHLIVQVSPKIVEISDSISNEGNNISLT 156
QY 108 GMLTFYRKTPTVFWQVNVOSFNALIVYSNRSGDPTITRQLGTAIVYSATTGAVALTG 167
DB 157 -CIAT-GRPEPTVLT-WRIISPKAVGFVSEDEYLEIGITREQSGDYECSSNDVAAPVR 213
QY 168 LKSLTKHLPLV---GRFVPPAAVAANNC 193
DB 214 RYKVTNTPPYISBAKGVGVGQKGTLOC 243

Search completed: February 3, 2004, 09:16:01
Job time : 35 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2004, 09:11:48 ; Search time 41 Seconds

(without alignment)
1010.430 Million cell updates/sec

Title: US-09-990-415a-2

Perfect score: 1345
Sequence: 1 MESSNGERPLDINIQPRWD.....PMAIPLIMTLKDKFLK 261

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: A_genseq_19Jun03:*
2: /SIDSI/gcgdata/genseq/genseq-emb1/AA1980.DAT:*
3: /SIDSI/gcgdata/genseq/genseq-emb1/AA1981.DAT:*
4: /SIDSI/gcgdata/genseq/genseq-emb1/AA1982.DAT:*
5: /SIDSI/gcgdata/genseq/genseq-emb1/AA1983.DAT:*
6: /SIDSI/gcgdata/genseq/genseq-emb1/AA1984.DAT:*
7: /SIDSI/gcgdata/genseq/genseq-emb1/AA1985.DAT:*
8: /SIDSI/gcgdata/genseq/genseq-emb1/AA1986.DAT:*
9: /SIDSI/gcgdata/genseq/genseq-emb1/AA1987.DAT:*
10: /SIDSI/gcgdata/genseq/genseq-emb1/AA1988.DAT:*
11: /SIDSI/gcgdata/genseq/genseq-emb1/AA1989.DAT:*
12: /SIDSI/gcgdata/genseq/genseq-emb1/AA1990.DAT:*
13: /SIDSI/gcgdata/genseq/genseq-emb1/AA1991.DAT:*
14: /SIDSI/gcgdata/genseq/genseq-emb1/AA1992.DAT:*
15: /SIDSI/gcgdata/genseq/genseq-emb1/AA1993.DAT:*
16: /SIDSI/gcgdata/genseq/genseq-emb1/AA1994.DAT:*
17: /SIDSI/gcgdata/genseq/genseq-emb1/AA1995.DAT:*
18: /SIDSI/gcgdata/genseq/genseq-emb1/AA1996.DAT:*
19: /SIDSI/gcgdata/genseq/genseq-emb1/AA1997.DAT:*
20: /SIDSI/gcgdata/genseq/genseq-emb1/AA1998.DAT:*
21: /SIDSI/gcgdata/genseq/genseq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/genseq/genseq-emb1/AA2001.DAT:*
23: /SIDSI/gcgdata/genseq/genseq-emb1/AA2002.DAT:*
24: /SIDSI/gcgdata/genseq/genseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1345	100.0	261	ABB76445	Human protein clus
2	1345	100.0	266	ABB64302	Human ORF672, Hom
3	1345	100.0	312	ABB06151	Human NS protein s
4	1345	100.0	325	AAW93760	Human polypeptide,
5	1345	100.0	374	AAW39967	Human polypeptide
6	1345	100.0	374	AAW60083	Human transport pr
7	1269	94.3	251	AAW41589	Human ORF ORF1353
8	1255	93.3	397	ABG20175	Novel human diagno
9	1089	81.0	322	AAW40413	Human polypeptide

10	1089	81.0	322	23	ABB79572	Alzheimer's disease
11	1089	81.0	322	23	ABG68039	Human Alzheimer's
12	1089	81.0	322	23	ABB76448	Human protein clus
13	1089	81.0	322	21	AAE13276	Human transporters
14	1089	81.0	332	21	AAW41585	Human ORF ORF1349
15	1089	81.0	351	22	AAW42198	Human polypeptide
16	1082	80.4	322	22	AAW62389	Rat tricarboxylate
17	977	72.6	325	23	AAU91141	Human secreted pro
18	975	72.5	230	23	AAU91139	Human secreted pro
19	975	72.5	251	23	AAU91140	Human secreted pro
20	876	65.1	261	23	ABW79573	Alzheimer's disease
21	876	65.1	261	23	ABW68040	Amino acid sequenc
22	851	63.3	242	22	AAW93934	Human protein sequ
23	754	56.1	321	22	ABB65884	Human protein sequ
24	753	56.0	322	22	ABB76447	Drosophila melanog
25	750	55.8	420	22	AAW40491	Human protein clus
26	749	55.7	322	22	AAW38705	Human polypeptide
27	749	55.7	322	23	AAE21177	Human polypeptide
28	726	54.0	236	22	AAW95218	Human TRICH-21 pro
29	636.5	47.3	195	22	AAU32503	Human protein sequ
30	602	44.8	174	22	AAU19596	Novel human secret
31	593.5	44.1	351	22	ABB62710	Human diagnostic a
32	545	40.5	132	22	ABG20174	Drosophila melanog
33	545	40.5	203	23	ABB97807	Novel human diagno
34	503	37.4	112	23	ABB79571	Human secretory po
35	480.5	35.7	340	23	AAE22908	Alzheimer's disease
36	368.5	27.4	77	23	ABB76446	Human transporter
37	312	23.2	127	23	ABB90092	Human protein clus
38	301	22.4	138	22	AAW41753	Human polypeptide
39	258	19.2	98	21	AAW03234	Human secreted pro
40	143.5	10.7	337	23	ABG61531	Human transporter
41	143.5	10.7	352	22	AAW25384	Human protein sequ
42	143	10.6	54	22	AAW76166	Human colon cancer
43	139	10.3	305	23	ABW06187	Human zinc-finger
44	117	8.7	79	21	AAW02779	Human secreted pro
45	96	7.1	775	22	AAU02904	Angiotensin conver

ALIGNMENTS

RESULT 1
ID ABB76445 standard; Protein; 261 AA.
XX ABB76445;
AC ABB76445;
XX 02-SEP-2002 (first entry)
XX
XX Human protein cluster I polypeptide.
DE
XX Human protein cluster I; human; metabolic disorder; obesity; diabetes;
XX Protein cluster I; human; metabolic disorder; obesity; diabetes;
KW antidiabetic; diagnosis; therapy.
XX
XX Homo sapiens.
OS
XX WO200242324-A1.
PN
XX 30-MAY-2002.
PD
XX 22-NOV-2001; 2001WO-SE02581.
PF
XX 24-NOV-2000; 2000SE-0004325.
PR
XX (PMAA) PHARMACIA AB.
PA
XX Altersand A;
PI
XX WPI; 2002-500277/53.
XX N-PSDB; ABB83754.
DR
XX Novel nucleic acid molecule encoding Protein Cluster I, useful in the
PT diagnosis of metabolic diseases, such as obesity and diabetes, and in

PT the identification of agents useful in the treatment of the diseases
XX
PS Claim 3; Page 24-25; 34pp; English.
XX
CC The present invention relates to the identification of a human gene
CC family (see ABN83734-57) encoding a group of polypeptides (see
CC ABN76445-48) referred to as Protein Cluster I. This family of
CC homologous proteins was identified by an 'all-versus-all' BLAST
CC procedure using all Caenorhabditis elegans proteins in the
CC Wormpep20 database release. Proteins of unknown function were
CC compared to the Drosophila melanogaster Flybase database, and
CC non-annotated protein clusters, conserved in both C. elegans and D.
CC melanogaster, were used in a BLAST procedure against the Celera
CC Human Genome Database, and Protein Cluster I proteins of unknown
CC function were selected for study. The human part of Protein
CC Cluster I comprises polypeptides encoded by 3 genes. The partial
CC gene sequence encoding the present polypeptide is expressed
CC primarily in the nervous system and digestive system. The claimed
CC protein cluster I nucleic acid molecules and proteins are proposed
CC to be useful for differential identification of the tissue(s) or
CC cell type(s) present in a biological sample, for diagnosis of
CC diseases and disorders, including metabolic disorders and immune
CC disorders, especially obesity and diabetes, and for identifying
CC agents useful in the treatment of such diseases.
XX
SQ Sequence 261 AA;
XX
Query Match 100.0%; Score 1345; DB 23; Length 261;
Best Local Similarity 100.0%; Pred. No. 2,1e-144;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MESKMGELPDLINIQEPRMDSTFLGRARHFFVTYDPRNLLSGAQLASRNIVQNYRAG 60
DB 1 MESKMGELPDLINIQEPRMDSTFLGRARHFFVTYDPRNLLSGAQLASRNIVQNYRAG 60
QY 61 VVTGGITEDQIMRAKYVYDSAFHPDGTGEKVVLLGRMSAQPVMNTTIGCMLEFRKPTV 120
DB 61 VVTGGITEDQIMRAKYVYDSAFHPDGTGEKVVLLGRMSAQPVMNTTIGCMLEFRKPTV 120
QY 121 VFQMVNQSFNALVYNSRSGDPTITVROLGTAAYVATGAVATLGLKSLTKLPLPLVG 180
DB 121 VFQMVNQSFNALVYNSRSGDPTITVROLGTAAYVATGAVATLGLKSLTKLPLPLVG 180
QY 181 RFVFPFAVAANNCINIPLMRQRELQVGI PVADDEAGRLGYSVTAKQGI FQVVISRICMA 240
DB 181 RFVFPFAVAANNCINIPLMRQRELQVGI PVADDEAGRLGYSVTAKQGI FQVVISRICMA 240
QY 241 IPMAIIPPLIMDTLEKDFLK 261
DB 241 IPMAIIPPLIMDTLEKDFLK 261
XX
RESULT 2
ID ABP64302 standard; Protein: 266 AA.
AC ABP64302;
XX
DT 04-NOV-2002 (first entry)
XX
DB Human ORF672.
XX
KW Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnary;
KW Antiflammatory; gene therapy; human; ORFX; atherogenic; platelet;
KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;
KW cancer; cardiovascular disease; allergy; autoimmune disease;
KW wound healing; blood coagulation disorder; inflammatory disorder.
XX
OS Homo sapiens.
XX
XX US2002082206-A1.
XX
XX 27-JUN-2002.

XX
PF 30-MAY-2001; 2001US-0867550.
XX
PR 30-MAY-2000; 2000US-208427P.
XX
PA (LEACH/) LEACH M D.
XX
PA (MEHR/) MEHRABAN F.
XX
PA (CONL/) CONLEY P B.
XX
PA (TOPP/) TOPPER J N.
XX
PA (LAWD/) LAW D.
XX
PI Leach MD, Mehraban F, Conley PB, Topper JN, Law D;
XX
DR WPI: 2002-626554/67;
XX
DR N-PSDB; ABQ98865.
XX
PT New polypeptide designated ORFX are present in human atherogenic cells
XX
PT and are useful to prevent and treat ORFX-associated disorders including
XX
PT cancer, allergy, wound healing or autoimmune, cardiovascular or
XX
PT inflammatory disease -
XX
PS Claim 10; SEQ ID 1344; 78pp; English.
XX
CC The present invention relates to novel human ORFX polypeptides and their
XX
CC coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences
XX
CC were discovered in human atherogenic cells, in particular in platelets
XX
CC and human umbilical vein endothelial cells (HUVEC) and are expressed in
XX
CC many other tissues as well. Atherogenic cells are cells which have the
XX
CC potential to develop atherosclerotic plaques. The ORFX polypeptides and
XX
CC nucleic acids are useful for treating or preventing a pathological
XX
CC condition associated with an ORFX-associated disease, e.g. cancer, blood
XX
CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood
XX
CC coagulation disorders or inflammatory disorders.
XX
CC Note: The sequence data for this patent did not form part of the printed
XX
CC specification, but was obtained in electronic format directly from the
XX
CC USPTO web site at seqdata.uspto.gov/sequence.html?docid=20020082206.
XX
SQ Sequence 266 AA;
XX
Query Match 100.0%; Score 1345; DB 23; Length 266;
Best Local Similarity 100.0%; Pred. No. 2,2e-144;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MESKMGELPDLINIQEPRMDSTFLGRARHFFVTYDPRNLLSGAQLASRNIVQNYRAG 60
DB 1 MESKMGELPDLINIQEPRMDSTFLGRARHFFVTYDPRNLLSGAQLASRNIVQNYRAG 60
QY 61 VVTGGITEDQIMRAKYVYDSAFHPDGTGEKVVLLGRMSAQPVMNTTIGCMLEFRKPTV 120
DB 61 VVTGGITEDQIMRAKYVYDSAFHPDGTGEKVVLLGRMSAQPVMNTTIGCMLEFRKPTV 120
QY 121 VFQMVNQSFNALVYNSRSGDPTITVROLGTAAYVATGAVATLGLKSLTKLPLPLVG 180
DB 121 VFQMVNQSFNALVYNSRSGDPTITVROLGTAAYVATGAVATLGLKSLTKLPLPLVG 180
QY 181 RFVFPFAVAANNCINIPLMRQRELQVGI PVADDEAGRLGYSVTAKQGI FQVVISRICMA 240
DB 181 RFVFPFAVAANNCINIPLMRQRELQVGI PVADDEAGRLGYSVTAKQGI FQVVISRICMA 240
QY 241 IPMAIIPPLIMDTLEKDFLK 261
DB 241 IPMAIIPPLIMDTLEKDFLK 261
XX
RESULT 3
ID ABB06151 standard; Protein: 312 AA.
AC ABB06151;
XX
XX 10-MAY-2002 (first entry)
XX
XX Human NS protein sequence SEQ ID NO:243.
DB

XX Human: cyrostatic; osteopathic; gynaecological; neuroprotective;
 KM antineumatic; antiaerthetic; antipsoeptic; ophthalmological; anti-HIV;
 KM vasectropic; antiarteriosclerotic; antinephrotoxic; dermatological;
 KM anorectic; muscular; antifertility; cardiovascular; anticoagulant;
 KM antifibrinolytic; hypotension; antiaesthetic; immunomodulator; cardiac;
 KM anticonvulsant; antidiabetic; tranquilliser; antidepressant; neuroleptic;
 KM gastrointestinal; virucide; antitumor; cerebroprotective; noctropic;
 KM contraceptive; vaccine; gene therapy; cancer; osteoporosis; dystonia;
 KM endometriosis; degenerative disease; multiple sclerosis; psoriasis;
 KM rheumatoid arthritis; cataract; restenosis; atherosclerosis; glaucoma;
 KM inflammation; skin disorder; obesity; muscular dystrophy; AIDS;
 KM infertility; cardiovascular disease; coagulation disease; hypertension;
 KM ischaemia; asthma; immune disease; epilepsy; angina; neurodegeneration;
 KM diabetes; anxiety; depression; schizophrenia; viral disease; stroke;
 KM gastric ulcer; Alzheimer's disease.

XX Homo sapiens.
 OS WO200206315-A2.
 PN 24-JAN-2002.
 PD 17-JUL-2001; 2001WO-IL00653.
 PF 18-JUL-2000; 2000IL-0137345.
 PR 15-DEC-2000; 2000IL-0140354.
 XX (COMP-) COMPUGEN LTD.
 PA Maintz L, Freilich S, Bernstein J;
 PI WPI; 2002-155037/20.
 DR N-PSDB; ABL39805.

XX One hundred and twenty eight novel nucleic acid sequences, useful for
 PT treating and diagnosing e.g. cancer, asthma and Alzheimer's -
 PT
 PS Claim 6; Page 278-279; 290pp; English.

XX ABL39691 to ABL39818 represent novel human nucleic acid sequences
 CC encoding the proteins given in ABL39691 to ABL39818. The novel sequences
 CC (NS) can have cyrostatic, osteopathic, gynaecological, neuroprotective,
 CC antineumatic, antiaerthetic, antipsoeptic, ophthalmological, virucide,
 CC vasectropic, antiarteriosclerotic, antinephrotoxic, dermatological,
 CC anorectic, muscular, antifertility, hypotension, cardiovascular, cardiac,
 CC anticonvulsant, antidiabetic, tranquilliser, antidepressant, neuroleptic,
 CC immunomodulator, anticonvulsant, antidiabetic, tranquilliser, antitumor,
 CC antidepressant, gastrointestinal, neuroleptic, cerebroprotective,
 CC contraceptive and contraceptive activities. The NS can be used in vaccines,
 CC gene therapy and antisense therapy. Nucleic acids, expression vectors and
 CC antibodies from the present invention can be used for treating and
 CC diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative
 CC diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,
 CC cataracts, restenosis, atherosclerosis, inflammation, skin disorders,
 CC glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
 CC disease, coagulation disease, ischaemia, hypertension, asthma, immune
 CC disease, epilepsy, angina, neurodegeneration, diabetes, anxiety, immune
 CC depression, schizophrenia, viral disease, gastric ulcers, stroke,
 CC Alzheimer's disease and as a contraceptive.

XX Sequence 312 AA;
 SQ
 Query Match 100.0%; Score 1345; DB 23; Length 312;
 Best Local Similarity 100.0%; Pred. No. 2.8e-144;
 Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESKMGELPLDINIOEPWDOSTFLGRARHFTYDPRNLLSGAQLASNNIVONRAG 60
 DB 1 MESKMGELPLDINIOEPWDOSTFLGRARHFTYDPRNLLSGAQLASNNIVONRAG 60
 QY 61 VTFPGITEDQLMRKYYVDSAFHPDTGKVKVLIGRMSAQVPMNMNTITGCMLEFRKPTV 120

DB 61 VTFPGITEDQLMRKYYVDSAFHPDTGKVKVLIGRMSAQVPMNMNTITGCMLEFRKPTV 120
 QY 121 VFWQWNVQSFPAIVYNSNRSGDPTITRQIGTAYVATGVAAPALGLKSLTKLPLVIG 180
 DB 121 VFWQWNVQSFPAIVYNSNRSGDPTITRQIGTAYVATGVAAPALGLKSLTKLPLVIG 180
 QY 181 RFVFFAFAVAAANCINIFLMRQRELOVGI PVADKAGORLGSVTPAKOGIFPVVISRICMA 240
 DB 181 RFVFFAFAVAAANCINIFLMRQRELOVGI PVADKAGORLGSVTPAKOGIFPVVISRICMA 240
 QY 241 IPMAVPEPLIMDTLEKDFLK 261
 DB 241 IPMAVPEPLIMDTLEKDFLK 261

RESULT 4
 ID AAM93760 standard; Protein; 325 AA.
 XX AAM93760;
 AC AAM93760;
 DT 06-NOV-2001 (first entry)
 DE Human polypeptide, SEQ ID NO: 3752.
 XX Human; full length cDNA; cDNA synthesis; oligo-capping.
 OS Homo sapiens.
 XX EP1130094-A2.
 PN 05-SEP-2001.
 PD 07-JUL-2000; 2000EP-0114089.
 PF 08-JUL-1999; 99JP-0194486.
 PR 11-JAN-2000; 2000JP-0118774.
 PR 02-MAY-2000; 2000JP-0183765.
 XX (HELI-) HELIX RES INST.
 PA Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 DR WPI; 2001-524255/58.
 N-PSDB; AAK94713.

XX 830 Primers useful for synthesizing full length cDNA clones and their
 PT use in genetic manipulation -
 PT
 PS Claim 8; SEQ ID NO 3752; 1380pp + sequence listing; English.

XX The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been
 CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
 CC molecules have been determined. Primers for synthesizing the full length
 CC cDNA are useful for clarifying the function of the protein encoded by
 CC the cDNA. The full length clones were obtained by construction of full
 CC length enriched cDNA libraries that were synthesized by the oligo-capping
 CC method. The primers enable the production of the full length cDNA easily
 CC without any special methods. The present sequence is a polypeptide
 CC encoded by a full length human cDNA of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in CD-ROM format directly from EPO.

XX Sequence 325 AA;
 SQ
 Query Match 100.0%; Score 1345; DB 22; Length 325;
 Best Local Similarity 100.0%; Pred. No. 3e-144;
 Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESKMGELPLDINIOEPWDOSTFLGRARHFTYDPRNLLSGAQLASNNIVONRAG 60

Db 1 MESKGEPLPDINIQEPMDQSTFLGRARHFTVTDPRNLLSGAQLEASRNIVQNRAG 60
 QY VTPGTTEDQMPRAKYVDSAFHPDTEGEKVLIGRMSAQVPMNTTTCMLTFYRKPTV 120
 Db 61 VTPGTTEDQMPRAKYVDSAFHPDTEGEKVLIGRMSAQVPMNTTTCMLTFYRKPTV 120
 QY 121 VFWQWVQSFNAIVNYSNRSGDPTITVRQLGTAVYSATGAVATPAGLKSITKHLPLV 180
 Db 121 VFWQWVQSFNAIVNYSNRSGDPTITVRQLGTAVYSATGAVATPAGLKSITKHLPLV 180
 QY 181 RFPVFAVAANANCINIPMRQRELQVGI PVADENGRIGSVTPRAKQGIQVWISRCMA 240
 Db 181 RFPVFAVAANANCINIPMRQRELQVGI PVADENGRIGSVTPRAKQGIQVWISRCMA 240
 QY 241 IPMAAIPPLIMDTLEKKDFLK 261
 Db 241 IPMAAIPPLIMDTLEKKDFLK 261

RESULT 5
 AAM39967
 ID AAM39967 standard; Protein: 374 AA.

AC AAM39967;

DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 3112.

XX Human, noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.

OS Homo sapiens.

PN WO200153312-A1.

PD 26-JUL-2001.

PF 26-DEC-2000; 2000WO-US34263.

PR 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.

PR 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

PA Tang YT, Liu C, Aaround V, Chen R, Ma Y, Qian XB, Ren F, Wang D,

PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

PI Zhao QJ, Zhou P, Goodrich R, Drmanac RT;

XX WPI: 2001-442253/47.

DR N-PSDB; AAI59123.

XX Novel nucleic acids and polypeptides, useful for treating disorders

XX such as central nervous system injuries -

XX Example 4; SEQ ID NO 3112; 10078pp; English.

CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: immune system suppression,
 CC activation/inhibition activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 SQ Sequence 374 AA;
 QY Query Match 100.0%; Score 1345; DB 22; Length 374;
 Db Best Local Similarity 100.0%; Pred. No. 3.7e-144;
 QY Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1 MESKGEPLPDINIQEPMDQSTFLGRARHFTVTDPRNLLSGAQLEASRNIVQNRAG 60
 QY VTPGTTEDQMPRAKYVDSAFHPDTEGEKVLIGRMSAQVPMNTTTCMLTFYRKPTV 120
 Db 61 VTPGTTEDQMPRAKYVDSAFHPDTEGEKVLIGRMSAQVPMNTTTCMLTFYRKPTV 120
 QY 121 VFWQWVQSFNAIVNYSNRSGDPTITVRQLGTAVYSATGAVATPAGLKSITKHLPLV 180
 Db 121 VFWQWVQSFNAIVNYSNRSGDPTITVRQLGTAVYSATGAVATPAGLKSITKHLPLV 180
 QY 181 RFPVFAVAANANCINIPMRQRELQVGI PVADENGRIGSVTPRAKQGIQVWISRCMA 240
 Db 181 RFPVFAVAANANCINIPMRQRELQVGI PVADENGRIGSVTPRAKQGIQVWISRCMA 240
 QY 241 IPMAAIPPLIMDTLEKKDFLK 261
 Db 241 IPMAAIPPLIMDTLEKKDFLK 261

RESULT 6
 AAB60083
 ID AAB60083 standard; Protein: 374 AA.

AC AAB60083;

DT 28-MAR-2001 (first entry)

DE Human transport protein TPPT-3.

XX Human, transport protein; TPPT; transport disorder; metabolic disorder;

KW neurological disorder; cardiovascular disorder; reproductive disorder;

KW immune disorder; cancer.

OS Homo sapiens.

PN WO200078953-A2.

PD 28-DEC-2000.

PF 16-JUN-2000; 2000WO-US16668.

PR 17-JUN-1999; 99US-0139923.

PR 10-AUG-1999; 99US-0148177.

PR 18-AUG-1999; 99US-0149357.

PR 28-OCT-1999; 99US-0162287.

XX (INCY-) INCYTE GENOMICS INC.

PA Lal P, Yang J, Yue H, Hillman JL, Tang YT, Bandman O, Burford N;

PI Baughn MR, Azimzai Y, Lu DM, Au-Young J, Patterson C;

XX WPI: 2001-041424/05.

DR N-PSDB; AAF27703.

PT Isolated polypeptide with a human transport protein sequence is useful
 PT for the diagnosis, prevention and treatment of disorders associated
 PT with the immune, reproductive and cardiovascular systems -

XX Claim 2; Page 108-109; 165pp; English.

XX The present invention provides the protein and coding sequences for 43
 CC novel human transport proteins (designated TPTs). These can be used in
 CC the diagnosis and treatment of transport, metabolic, neurological,
 CC reproductive, cardiovascular and immune disorders, and cell proliferative
 CC disorders such as cancer.

XX Sequence 374 AA;

Query Match 100.0%; Score 1345; DB 22; Length 374;

Best Local Similarity 100.0%; Pred. No. 3.7e-144;

Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESKNGELPLDINIOEPRMDSTFLGRARHFTVTDPRNLLLSGAQLEASRNIVONYRAG 60
 DB 1 MESKNGELPLDINIOEPRMDSTFLGRARHFTVTDPRNLLLSGAQLEASRNIVONYRAG 60
 QY 61 VVTPEITDOLMRARVYVDSAFHPDTEKRVLLIGMSAQVPMNMITTCMLTFYKRTPTV 120
 DB 61 VVTPEITDOLMRARVYVDSAFHPDTEKRVLLIGMSAQVPMNMITTCMLTFYKRTPTV 120
 QY 121 VFWOVNOSFNAIVVNSRSGDPTITVROLGTAIVYSAATGAVALTGLSKLTKEHLPVLV 180
 DB 121 VFWOVNOSFNAIVVNSRSGDPTITVROLGTAIVYSAATGAVALTGLSKLTKEHLPVLV 180
 QY 181 RVFPFAVAANCINIPLMRQRELVGIVADAGORLGYSYTAAGKQIFQVVISRICMA 240
 DB 181 RVFPFAVAANCINIPLMRQRELVGIVADAGORLGYSYTAAGKQIFQVVISRICMA 240
 QY 241 IPMAIIPPLIMDTLEKDFLK 261
 DB 241 IPMAIIPPLIMDTLEKDFLK 261

RESULT 7
 AAB41589

ID AAB41589 standard; Protein; 251 AA.

AC AAB41589;

DT 08-FEB-2001 (first entry)

Human ORFX ORF1353 polypeptide sequence SEQ ID NO:2706.

XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 XX vulnerability; antiproliferative; antiparkinsonian; neurotropic; neuroprotective;
 XX anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 XX immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 XX hypotensive; dermatological; immunosuppressive; antineoplastic; antitumor;
 XX antiviral; antibacterial; antifungal; antipneumatic; antithyroid;
 XX antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 XX neurodegenerative disorder; osteoarthritis; graft vs host disease;
 XX cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 XX cholesterol ester storage; systemic lupus erythematosus; infection;
 XX severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 XX allergy; aplastic anemia; nocturnal haemoglobinuria; burn; wound;
 XX bone damage; cartilage damage; nocturnal haemoglobinuria; burn; wound;
 XX thrombosis; contraceptive.

XX Homo sapiens.

XX WO200058473-A2.

XX 05-OCT-2000.

XX 31-MAR-2000; 2000WO-US08621.

XX 31-MAR-1999; 99US-0127607.

PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.

XX (CURA-) CURAGEN CORP.

XX Shinketsu RA, Leach M;

XX WPI, 2000-602362/57.

XX N-PSDB; AAC75798.

PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -

XX Claim 11; Page 1942-1943; 5507pp; English.

CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnerability;
 CC antiproliferative; antiparkinsonian; neurotropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antineoplastic; antibacterial; antiviral; antifungal; antipneumatic;
 CC antithyroid; and antianemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancer,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antineoplastic diseases, to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.

XX Sequence 251 AA;

Query Match 94.3%; Score 1269; DB 21; Length 251;

Best Local Similarity 98.0%; Pred. No. 9.2e-136;

Matches 246; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 MESKNGELPLDINIOEPRMDSTFLGRARHFTVTDPRNLLLSGAQLEASRNIVONYRAG 60
 DB 1 MESKNGELPLDINIOEPRMDSTFLGRARHFTVTDPRNLLLSGAQLEASRNIVONYRAG 60
 QY 61 VVTPEITDOLMRARVYVDSAFHPDTEKRVLLIGMSAQVPMNMITTCMLTFYKRTPTV 120
 DB 61 VVTPEITDOLMRARVYVDSAFHPDTEKRVLLIGMSAQVPMNMITTCMLTFYKRTPTV 120
 QY 121 VFWOVNOSFNAIVVNSRSGDPTITVROLGTAIVYSAATGAVALTGLSKLTKEHLPVLV 180
 DB 121 VFWOVNOSFNAIVVNSRSGDPTITVROLGTAIVYSAATGAVALTGLSKLTKEHLPVLV 180
 QY 181 RVFPFAVAANCINIPLMRQRELVGIVADAGORLGYSYTAAGKQIFQVVISRICMA 240
 DB 181 RVFPFAVAANCINIPLMRQRELVGIVADAGORLGYSYTAAGKQIFQVVISRICMA 240
 QY 241 IPMAIIPPLIM 251
 DB 241 IPMAIIPPLIM 251

RESULT 8

ID AAG20175 standard; Protein; 397 AA.

AC AAG20175;

DT 18-FEB-2002 (first entry)

```

XX DE Novel human diagnostic protein #20166.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW Human; supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PE 30-MAR-2001; 2001WO-US08631.
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX PA (HYSE-) HYSEQ INC.
XX P1 Drmanac RT, Liu C, Tang YT;
XX DR WPI; 2001-639362/73.
XX DR N-PSDB; AAS84362.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensic, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.
XX PS Claim 20; SEQ ID NO 50534; 103bp; English.
XX CC The invention relates to isolated polynucleotide (I) and
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX CC and gene mapping, and in recombinant production of (II). The
XX CC polynucleotides are also used in diagnostics as expressed sequence tags
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques
XX CC to restore normal activity of (II) or to treat disease states involving
XX CC (II). (II) is useful for generating antibodies against it, detecting or
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
XX CC a food supplement. (II) and its binding partners are useful in medical
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
XX CC disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensic, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. ABG00010-ABG0377 represent novel human
XX CC diagnostic amino acid sequences of the invention.
XX CC Note: The sequence data for this patent did not appear in the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 397 AA;

```

Query Match 93.3%; Score 1255; DB 22; Length 397;
 Best Local Similarity 89.6%; Pred. No. 7,4e-134;
 Matches 251; Conservative 0; Mismatches 7; Indels 22; Gaps 2;

```

Oy 4 KMGELPLDINIOEPWOSTFLGRARHPTTDPNNLLSGAOLASNNIYONRAGVVT 63
Db 70 KMGELPLDINIOEPWOSTFLGRARHPTTDPNNLLSGAOLASNNIYONRAGVVT 129
Oy 64 PGITEDQLMRAKYVYDSAFHPTGKVVLLIGMSAQVPMNNTITGCMLEFY----- 114
Db 130 PGITEDQLMRAKYVYDSAFHPTGKVVLLIGMSAQVPMNNTITGCMLEFYROGSKDEGH 189
Oy 115 -----RKPTTVFMQWUNOSRNATVYNRS-GPTPTVRLGLNAYVATGA 161
Db 190 CRRGRSECLSLRKTPTVFMQWUNOSRNATVYNRSKHPSTVRLGLNAYVATGA 249
Oy 162 VATAIGLSLTKHLPLVGRFVPPFAVAANCINIPLMRQRELQVIGIVADAGORLGYS 221

```

```

Db 250 VATAIGLSLTKHLPLVGRFVPPFAVAANCINIPLMRQRELQVIGIVADAGORLGYS 309
Oy 222 VTAAGQIFQVVISRICMAIPMAIPLIMDTLEKDKPLK 261
Db 310 VTAAGQIFQVVISRICMAIPMAIPLIMDTLEKDKPLK 349

```

RESULT 9
 AAM40413
 ID AAM40413 standard; Protein; 322 AA.
 AC AAM40413;
 XX 22-OCT-2001 (first entry)
 DE Human polypeptide SEQ ID NO 3558.
 XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX OS Homo sapiens.
 XX PN WO200153312-A1.
 XX PD 26-JUL-2001.
 XX PE 26-DEC-2000; 2000WO-US4263.
 XX PR 21-JAN-2000; 2000US-0488725.
 XX PR 25-APR-2000; 2000US-0552317.
 XX PR 09-JUL-2000; 2000US-0598042.
 XX PR 19-JUL-2000; 2000US-0620312.
 XX PR 03-AUG-2000; 2000US-0653450.
 XX PR 14-SEP-2000; 2000US-0662191.
 XX PR 19-OCT-2000; 2000US-0693036.
 XX PR 29-NOV-2000; 2000US-0727344.
 XX PA (HYSE-) HYSEQ INC.
 XX P1 Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 XX P1 Wang J, Wang Z, Wehrman T, Xu C, Xue AD, Yang Y, Zhang J;
 XX P1 Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX WPI; 2001-442253/47.
 XX N-PSDB; AA159569.
 XX PT Novel nucleic acids and polypeptides, useful for treating disorders
 XX PT such as central nervous system injuries -
 XX Example 8; SEQ ID NO 3558; 10078bp; English.
 XX The invention relates to human nucleic acids (AA157798-AA161369) and
 XX the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 XX immunosuppressant and cytostatic activity. The polynucleotides are useful
 XX in gene therapy. A composition containing a polypeptide or polynucleotide
 XX of the invention may be used to treat diseases of the peripheral nervous
 XX system, such as peripheral nervous injuries, peripheral neuropathy and
 XX localised neuropathies and central nervous system diseases, such as
 XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 XX utilisation of the activities such as: Immune system suppression,
 XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 XX assays for receptor activity, arthritis and inflammation, leukaemias and
 XX C.N.S disorders.
 XX Note: The sequence data for this patent did not form part of the printed
 XX specification.
 XX SQ Sequence 322 AA;

ABG68039
 ID ABG68039 standard; Protein; 322 AA.
 AC ABG68039;
 DT 07-OCT-2002 (first entry)
 DE Human Alzheimer's disease-associated protein isoform-41 (ADPI-41).
 XX Human; Alzheimer's disease; AD; brain tissue; ADP; ADPI;
 KM Alzheimer's disease-associated feature; neuroprotective;
 KM Alzheimer's disease-associated protein isoform; nootropic;
 KM ADPI-41.
 XX Homo sapiens.
 OS
 XX WO200246767-A2.
 PN 13-JUN-2002.
 XX
 PD 29-NOV-2001; 2001WO-GB05289.
 PF 08-DEC-2000; 2000US-254431P.
 PR (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
 PA Herath HMAC, Parekh RB, Rohlf C;
 PI WPI; 2002-508575/54.
 DR N-PSDB; ABK96715.
 XX
 XX Screening, diagnosis or prognosis of Alzheimer's disease in subject,
 PT comprises detecting Alzheimer's disease-associated features or
 PT Alzheimer's disease-associated protein isoforms in brain tissue
 PT from the subject -
 XX
 XX Example 2; Fig 2b; 427pp; English.
 PS The present invention relates to methods and compositions for the
 XX screening, diagnosis or prognosis of Alzheimer's disease (AD) in
 XX a subject. The method comprises analysing a sample of brain tissue
 CC from a subject by 2D electrophoresis to generate a 2D array of
 CC Alzheimer's disease-associated features (ADFs), whose relative
 CC abundance correlates with the presence, absence, stage or severity of
 CC AD and comparing the abundance of each feature with the abundance of
 CC that chosen feature in brain tissue from persons free from AD. The
 CC invention also describes Alzheimer's disease-associated protein
 CC isoforms (ADPIs) detectable in brain tissue. The methods and
 CC compositions of the invention are useful for the screening, diagnosis
 CC or prognosis of AD in a subject, for determining the stage or severity
 CC of AD in a subject, for identifying a subject at risk of developing AD,
 CC or for monitoring the effect of therapy administered to a subject
 CC having AD. Antibodies capable of binding to ADPIs are useful for
 CC treating or preventing AD, and for determining the efficacy of a given
 CC treatment regime. An agent that modulates the activity of ADPI is
 CC useful in the manufacture of a medicament for the treatment or
 CC prevention of AD in a subject. The present sequence represents human
 CC ADPI-41.
 XX
 SQ Sequence 322 AA;
 Query Match 81.0%; Score 1089; DB 23; Length 322;
 Best Local Similarity 78.5%; Pred. No. 4,4e-115;
 Matches 201; Conservative 28; Mismatches 27; Indels 0; Gaps 0;
 QY 6 GELPDLINIOEPKRDOSTFLGRARHFFVTTPRNLLISGAQLSEARNIVQYRAGVYTPG 65
 DB 3 GELPNNINIKPRKPDOSTFIRGRANHFVTTPRNILLTNEDESAKIVHDIYRGGIYPPG 62
 QY 66 ITEDQLMRKXVYDASFPDGTGKRVLLIGRMSAQVPMNMITTCMLTFYRKTPTVFMQW 125
 DB 63 LTEMELWRKAKIYDASFPDGTGKRVLLIGRMSAQVPMNMITTCMLTFYRKTPTVFMQW 122

QY 126 VNOSFNAIVNYSRSGDPTITVROLGTAVYSATTGAATLGLKSLTKHLPVLGSRVPP 185
 DB 123 INOSFNAIVNYSRSGDPTITVROLGTAVYSATTGAATLGLKSLTKHLPVLGSRVPP 182
 QY 186 AAVAANNCINIPLMRQRELOVGIPIVADEAGQRLGYSVTAKQGIPOVVISRICMAIPAMA 245
 DB 183 AAVAANNCINIPLMRQRELOVGIPIVDENGNRIGESANAKKAITQVYVSRIMAAPGMA 242
 QY 246 IPPILMDTLEKKDFLK 261
 DB 243 IPPFIMNTLEKKAFLK 258
 RESULT 12
 ABB76448
 ID ABB76448 standard; Protein; 322 AA.
 AC ABB76448;
 DT 02-SEP-2002 (first entry)
 DE Human protein cluster I polypeptide.
 XX Protein cluster I; human; metabolic disorder; obesity; diabetes;
 KM antidiabetic; diagnosis; therapy.
 OS Homo sapiens.
 XX WO200242324-A1.
 PN 30-MAY-2002.
 XX
 PD 22-NOV-2001; 2001WO-SE02581.
 PF 24-NOV-2000; 2000SE-0004325.
 PR (PHAA) PHARMACIA AB.
 PA Alterband A;
 PI WPI; 2002-500277/53.
 DR N-PSDB; ABN83757.
 XX
 XX Novel nucleic acid molecule encoding Protein Cluster I, useful in the
 PT diagnosis of metabolic diseases, such as obesity and diabetes, and in
 PT the identification of agents useful in the treatment of the diseases
 XX
 PS Claim 3; Page 32-33; 34pp; English.
 XX The present invention relates to the identification of a human gene
 CC family (see ABN83754-57) encoding a group of polypeptides (see
 CC ABB76445-48) referred to as Protein Cluster I. This family of
 CC homologous proteins was identified by an 'all-versus-all' BLAST
 CC procedure using all Caenorhabditis elegans proteins in the
 CC Wormpep20 database release. Proteins of unknown function were
 CC compared to the Drosophila melanogaster flybase database, and
 CC non-annotated protein clusters, conserved in both C. elegans and D.
 CC melanogaster, were used in a BLAST procedure against the Celera
 CC Human Genome Database, and Protein Cluster I proteins of unknown
 CC function were selected for study. The human part of Protein
 CC Cluster I comprises polypeptides encoded by 3 genes. The gene
 CC encoding the present polypeptide is expressed primarily in the
 CC immune system. The claimed Protein Cluster I nucleic acid
 CC molecule and proteins are proposed to be useful for differential
 CC identification of the tissue(s) or cell type(s) present in a
 CC biological sample, for diagnosis of diseases and disorders,
 CC including metabolic disorders and immune disorders, especially
 CC obesity and diabetes, and for identifying agents useful in the
 CC treatment of such diseases.
 XX
 SQ Sequence 322 AA;
 Query Match 81.0%; Score 1089; DB 23; Length 322;

	Best Local Similarity	79.5%; Pred. No. 4.4e-115;	Mismatches 201; Conservative 28; Mismatches 27; Indels 0; Gaps 0
QY	6	GELPELDINIQEPNDQSTFTLGRARHFFVTVPDRNLLTSGAQLGASRNIVQNTAGVTPG	65
Db	3	GELPENNINIKERPDQSTFTIGRANHFVTVPDRNILLTNTEQLSARKIIVHDYQGIIVPG	62
QY	66	ITTEPOLBRAKVVYOSAFHPDGTGEKVLVLRMSAQVPMNMTTGCMLTFRKPTPYVFMQM	125
Db	63	LTELHELMRAKIITYSAFHPDTGEKKILLGRMSAQVPMNMTTTGCMHTYRTTRPAVLFMQM	122
QY	126	VNQSFNAIWVNSNRSGDPITVTRQLGTAVYSATTTGAVALGLSKLTJCHLPVLYGRPVDF	185
Db	123	INGSFNAVVMVNTNNSGDAPLVYNELGTAVYSATTGAVALGLNLMTGHVSPLIRFPDF	182
QY	186	AAVAANAACINIPLMRQRELQVGIPVADAGQRLGYSVTAAKGIIQOVVISRIQMAIPMA	245
Db	183	AAVAANAACINIPLMRQRELVKGIPTVDENGNRLGESANAQAQAIQQVVVSRIILMAAPGWA	242
QY	246	IPLIMDLTKKDDFLK 261	
Db	243	IPLPIIMTLKKKAFLEK 258	

OS	Homo sapiens.
XX	
PN	W0200177174-A2.
XX	
PD	18-OCT-2001.
XX	
PF	06-APR-2001; 2001MO-US11206.
XX	
PR	06-APR-2000; 2000US-195595P.
PR	12-APR-2000; 2000US-196872P.
PR	20-APR-2000; 2000US-199020P.
PR	28-APR-2000; 2000US-200552P.
PR	05-MAY-2000; 2000US-202346P.
PR	11-MAY-2000; 2000US-203495P.
PA	(INCYT-) INCYTE GENOMICS INC.
XX	
PI	Reddy R, Thornton M, Borowsky ML, Tang YT, Khan FA, Tribouley CK
PI	Gandhi AR, Yao MG, Sanjanwala MS, Baughn NK, Nguyen DB;
PI	Polichek JL, Yue H, Seifhamer JJ, Wallia NR, Lal P, Kearney L,
PI	Walsh RT, Lu DAM, Lu Y, Greene BD, Raumann BE, Patterson C;
XX	
DR	WPI; 2002-017448/02.
DR	N-PDB; AAD21995.
XX	
PT	Polypeptides of human transporters and ion channels, useful for
PT	diagnosing, treating or preventing disorders of transport,
XX	neurological, muscle, immunological and cell proliferative disorders
XX	Claim 1; Page 120; 150pp; English.

XX The invention relates to human transporters and ion channels (TRICH)
CC and the polynucleotides encoding them. The composition comprising TRICH
CC or agonist of TRICH is useful for treating a disease or condition
CC associated with decreased expression of functional TRICH or condition
CC associated with overexpression of TRICH respectively. The composition
CC comprising Ab is useful for diagnosing a condition of disease associated
CC with expression of TRICH in a subject, where the disorders include a
CC transport disorder such as akathisia, cystic fibrosis, diabetes mellitus,
CC Parkinson's disease, myasthenia gravis, cardiac disorders associated
CC with transport e.g. angina, hypertension, myocarditis, neurological
CC disorders associated with transport e.g. Alzheimer's disease, Wilson's
CC disease, schizophrenia, cataract, infertility, hypoglycaemia, Grave's
CC disease, goitre, Addison's disease, Huntington's disease, dementia,
CC multiple sclerosis, bacterial and viral meningitis. TRICH DNA is useful
CC for generating a transcript image of a tissue or cell type, which
CC represents the global pattern of gene expression by a particular tissue
CC or cell type and for analysing the proteome of a tissue or cell type.
CC TRICH DNA is used in gene therapy. The present amino acid sequence is
CC human TRICH3 protein.
XX
XX
XX Sequence 322 AA;
XQ

Query Match	81.0%;	Score 1089;	DB 23;	Length 322;
Best Local Similarity	78.5%;	Pred. 10.4e-115;		
Matches 201;	Conservative 28;	Mismatches 27;	Indels 0;	Gaps 0;
QY	6	GELPLDINIGEPWQDSTFLGARAHEFTVTDPRNLLISGAOLEASRNIVQNYRAGVYTPG	65	
Db	3	GELPPNINIKEPWQDSTFLGRANHFPTVTDPRILLTLTNEQLESARKIVHDYRQGIWPG	62	
QY	66	ITEDQLMRATYVYDSAFHPDPTGERVVLIGSMQAVPMNMITTGOMLFLPKRTPIVYVFWQ	125	
Db	63	LTENELMRARYKIDYSAHPDPTGERMTLIGSMQAVPMNMITTGOMFLPKRTPIVYVFWQ	122	
QY	126	VNOSFNATVYNSRSGDPTITVROLGTAAYVSATGAATATLGLKSLTKHLPLVLRGVFVF	185	
Db	123	INOSFNATVYNTSRSGAPLTVNELGTAAYVSATGAATATLGLNALTKHVSPLIGRVFVF	182	
QY	186	AAVAANNCINIPLMRORELOVGIPIVADCAQGLGYSVTAAKGIFQVYISRICHAIPMA	245	
Db	183	AAVAANNCINIPLMRORELVGIPVDENGNRIGESANAKQAITQVVSRIILMAAGMA	242	
QY	246	IPPLIMDTLEKDFLK 261		
Db	243	IPPLIMDTLEKDFLK 258		
RESULT 14				
AAB41585				
ID	AAB41585	standard; Protein: 332 AA.		
XX	AC	AAB41585;		
XX	DT	08-FEB-2001 (fast entry)		
XX	DE	Human ORFX ORF1349 polypeptide sequence SEQ ID NO:2698.		
XX	XX	Human; open reading frame; ORFX; detection; cyostatic; hepatotropic;		
KM	XX	vulneray; antipsoptic; antiparkinsonian; nootropic; neuroprotective;		
KM	XX	anticovulant; osteopathic; antipathritic; immunosuppressant; cardiac;		
KM	XX	immunosuppressant; thrombolytic; coagulant; vasotropic; antidiabetic;		
KM	XX	hypocensive; demagogical; immunosuppressive; antiinflammatory;		
KM	XX	antiviral; antibacterial; antifungal; antirheumatic; antithyroid;		
KM	XX	antidiabetic; gene therapy; cancer; proliferative disorder; hypertension;		
KM	XX	neurodegenerative disorder; osteoarthritis; graft vs host disease;		
KM	XX	cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;		
KM	XX	cholesterol ester storage; systemic lupus erythematosus; infection;		
KM	XX	severe combined immunodeficiency; malaria; autoimmune disorder; asthma;		
KM	XX	allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;		
KM	XX	bone damage; cartilage damage; antiinflammatory disease; coagulation;		
KM	XX	thrombosis; contraceptive.		

OS Homo sapiens.
 XX
 PN W0200058473-A2.
 XX
 PD 05-OCT-2000.
 XX
 PF 31-MAR-2000; 2000WO-US08621.
 XX
 PR 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shimkets RA, Leach M;
 XX
 DR WPI: 2000-602362/57.
 DR N-PSDB: AAC75794.
 XX
 PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 XX Claim 11; Page 1938; 5507pp; English.
 XX
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatocytotoxic; vulnery;
 CC antiproliferative; antiparasitism; neurotrophic; neuroprotective;
 CC osteoplastic; anticonvulsant; antidiabetic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasoconstrictive;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antidiabetic; antiparasitic; antiviral; antifungal; antirheumatic;
 CC antihypertensive; and antianemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, anti-inflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 CC
 XX Sequence 332 AA;
 SO
 Query Match 81.0%; Score 1089; DB 21; Length 332;
 Best Local Similarity 78.5%; Pred. No. 4.7e-115;
 Matches 201; Conservative 28; Mismatches 27; Indels 0; Gaps 0;
 QY 6 GELPDLINIOPRPDQSTFLGRARHFFVTTPRNLISGALQLEASRNIVQYRAGVTPRG 65
 DB 13 GELPDLINIKKPRPDQSTFLGRARHFFVTTPRNLISGALQLEASRNIVQYRAGVTPRG 72
 QY 66 IITPOLMRKAVVYVSAPFDTEGEKVLLIGRMSAQVPMAMTTTGCLTFYRKTPTVFWQW 125
 DB 73 LITBELMRKAVVYVSAPFDTEGEKVLLIGRMSAQVPMAMTTTGCLTFYRKTPTVFWQW 132
 QY 126 VNQSFNAIVNYSNRSGDTPTIVROLGTA VYSAATTGAVALGLSLTGLPLVGRFVPF 185
 DB 133 INQSFNAIVNYSNRSGDTPTIVROLGTA VYSAATTGAVALGLSLTGLPLVGRFVPF 192
 QY 186 AAVAAANCINTPLMRQRELOVIGIPVADAGQRLGVSVAACOGIFQVYVISTICMAIPAMA 245
 DB 193 AAVAAANCINTPLMRQRELOVIGIPVADAGQRLGVSVAACOGIFQVYVISTICMAIPAMA 252
 QY 246 IPPIMNTLEKKAFIX 261
 DB 253 IPPIMNTLEKKAFIX 268

RESULT 15
 AAM42198
 ID AAM42198 standard; Protein; 351 AA.
 XX
 AC AAM42198;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 7129.
 XX
 KW Human; neurotrophic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN W0200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US34263.
 XX
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang Z, Wehrman T, Xu C, Xue AD, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX
 DR WPI: 2001-442253/47.
 DR N-PSDB: AAI61354.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX
 XX Example 2; SEQ ID NO 7129; 10078pp; English.
 XX
 CC The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with neurotrophic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemia and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX
 XX Sequence 351 AA;
 SO
 Query Match 81.0%; Score 1089; DB 22; Length 351;
 Best Local Similarity 78.5%; Pred. No. 5.1e-115;
 Matches 201; Conservative 28; Mismatches 27; Indels 0; Gaps 0;

```
Qy 6 GELPLDINIOEPRMDSTFLGRARHPFTYTDPRNLLLSGQLEASRNIVQNYRAGVVTG 65
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 32 GELPRINIKERPMOSTFIRGRANHFVTDPNNILLTNEQLSARKIVHDIRQGI VPG 91
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 66 ITEDQIMRAKYVYDSAFHPDTGEKVVLIGRMSAQVPMNMITGCMLEFYRKPTPTVFMQM 125
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 92 LTENELMRAKYIYDSAFHPDTGEKMLIIRMSAQVPMNMITGCMWTFYRTTPAVLFWQM 151
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 126 VNGSFRAIVNGNRSGDTPIYRQLGTAYVSATTGAVATLGLKSLTKHLPLVGRFVPP 185
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 152 INQSFVAVVNYNTRSGDAPLTVELGTAYVSATTGAVATLGINALTKHVSPLIGRFVP 211
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 186 AAVAAANCINIPLMRORELOVGI PVADBAGORLGYSVTAAGIPOVVISRICMAIPAMA 245
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 212 AAVAAANCINIPLMRORELOVGI PVTDENGNRLGESANAKQAITOVVVSRIIMAAFGMA 271
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 246 IPPLIIMDTLEKDDFLK 261
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 272 IPPFIWNTLEKKAFLK 287
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

Search completed: February 3, 2004, 09:12:44
Job time : 42 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2004, 09:11:53 ; Search time 21 Seconds

(without alignments)
525.864 Million cell updates/sec

Title: US-09-990-415A-2

Perfect score: 1345

Sequence: 1 MESMGSLPLDINIOEPRWD.....PMAIPLIMTLKIOFLK 261

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: /cgm2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgm2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgm2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgm2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgm2_6/ptodata/1/1aa/PCTUS.COMB.pep:*
6: /cgm2_6/ptodata/1/1aa/Backfilest.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	96	7.1	602	2	US-08-419-652-6
2	96	7.1	783	6	Sequence 6, Appl1
3	96	7.1	836	1	Patent No. 5422248
4	89.5	6.7	863	1	US-07-923-976-4
5	89	6.6	655	4	US-09-134-001C-5531
6	88	6.5	394	4	US-09-252-991A-18502
7	82	6.1	146	2	US-08-627-173-20
8	82	6.1	146	2	US-08-535-882A-20
9	82	6.1	146	3	US-08-316-424A-6
10	82	6.1	146	3	US-09-005-546-20
11	82	6.1	146	4	US-08-477-669-6
12	82	6.1	771	1	US-07-923-976-6
13	81	6.0	514	4	US-09-252-991A-25845
14	81	6.0	676	3	US-08-947-965-71
15	80	5.9	621	4	US-09-252-991A-19749
16	80	5.9	815	4	US-09-914-259-18
17	79.5	5.9	529	5	PCT-US92-00282-7
18	79.5	5.9	693	1	US-08-463-620-11
19	79.5	5.9	693	2	US-08-224-917-11
20	79.5	5.9	693	2	US-08-914-853-11
21	79.5	5.9	693	2	PCT-US95-03934A-11
22	77	5.7	435	4	US-09-252-991A-17750
23	76	5.7	1209	4	US-09-252-991A-25844
24	75.5	5.6	531	5	PCT-US92-00282-5
25	75	5.6	562	4	US-09-773-466A-9
26	75	5.6	627	4	US-09-252-991A-23072
27	75	5.6	705	4	US-09-252-991A-27442

28	74	5.5	430	4	US-08-311-731A-153	Sequence 153, App
29	74	5.5	654	4	US-09-134-001C-3261	Sequence 3261, Ap
30	73.5	5.5	454	1	US-07-915-720D-17	Sequence 17, Appl
31	73.5	5.5	454	3	US-09-025-543-17	Sequence 17, Appl
32	73.5	5.5	456	1	US-07-915-720D-19	Sequence 19, Appl
33	73.5	5.5	456	3	US-08-879-565-11	Sequence 11, Appl
34	73.5	5.5	456	3	US-09-025-543-19	Sequence 19, Appl
35	73.5	5.5	456	3	US-08-431-517F-16	Sequence 16, Appl
36	73.5	5.5	487	1	US-08-030-644-2	Sequence 2, Appl1
37	73.5	5.5	487	1	US-08-013-801-2	Sequence 2, Appl1
38	73.5	5.5	487	1	US-08-072-063-2	Sequence 2, Appl1
39	73.5	5.5	487	1	US-08-212-132-2	Sequence 2, Appl1
40	73.5	5.5	487	1	US-08-414-924-2	Sequence 2, Appl1
41	73.5	5.5	487	1	US-08-311-611A-69	Sequence 69, Appl
42	73.5	5.5	487	1	US-08-311-611A-146	Sequence 146, App
43	73.5	5.5	487	1	US-08-173-968-2	Sequence 2, Appl1
44	73.5	5.5	487	1	US-08-232-527-2	Sequence 2, Appl1
45	73.5	5.5	487	1	US-08-372-783-69	Sequence 69, Appl

ALIGNMENTS

RESULT 1
US-08-419-652-6
Sequence 6, Application US/08419652
Patent No. 5831007
GENERAL INFORMATION:
APPLICANT: Chua, Anne O
APPLICANT: Gubler, Ulrich A
TITLE OF INVENTION: INTERLEUKIN-12 RECEPTOR
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESS: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: New Jersey
COUNTRY: United States of America
ZIP: 07110-1199
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/419,652
CLASSIFICATION: 530
FILING DATE: 11-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/248,532
FILING DATE: 31-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/094,713
FILING DATE: 19-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Kaes, Alan P
REGISTRATION NUMBER: 32142
REFERENCE/DOCKET NUMBER: CD 9174
TELECOMMUNICATION INFORMATION:
TELEPHONE: (201) 235-4205
TELEFAX: (201) 235-3500
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 602 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
NAME/KEY: Region
LOCATION: 1..602
OTHER INFORMATION: /note= "Represents residues 98 to
OTHER INFORMATION: 731 of human granulocyte colony-stimulating
OTHER INFORMATION: factor-receptor."

US-08-419-652-6

Query Match	7.1%;	Score 96;	DB 2;	Length 602;
Best Local Similarity	24.9%;	Pred. No. 0.028;		
Matches 54;	Conservative 34;	Mismatches 93;	Indels 36;	Gaps 12

```

Qy      61  VWPFGTEOOLMAK--YYUDSAFHDTEKEVL--IRMSQ--VPANNITIGMLTF 113
      Ds      384  IVTF-LYDPTMGSSQHVAYSQEMBSAPHELHGHGKTQAQLEWVEPPELOKSPETH 442
Qy      114  YRKTPTVFW-QWVNOSFNAIVNYSNR---GDTPTT--VRGLTAAYSATGAAYATA 165
      Ds      443  Y-----TIFWTAHQNSFSALINASRGVHLGLEPASVYHIIHMAAQAGATNSFVTLT 497
Qy      166  LGKLSLTKHLPLVGRF---VPPAAVAANCINIPMGQRELOVQIPADDAQGRIGSY 227
      Ds      498  MTLTPGSESLHIIIGLFGILLLTCTCGTAWLCCSPNRKNPLPMSVP--DEAHSLGSW 555
Qy      223  -TAAKGIFQVVISRICMAIPAMAPIPLIMDTLEKDD 258
      Ds      556  PTIMEADAFQ-----LPGLGTPTPIRKLVLVEED 583

```

RESULT 2
5422248-1

Patent No 5422248
 APPLICANT: SMITH, CRAIG A.; LARSEN, ALF D.; SIMS, JOHN E.;
 BENSON, CURTIS M.
 TITLE OF INVENTION: DNA SEQUENCES ENCODING GRANULOCYTE-COLONY
 STIMULATING FACTOR RECEPTORS
 NUMBER OF SEQUENCES: 6
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/6,183
 FILING DATE: 15-JAN-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 587,329
 FILING DATE: 24-SEP-1990
 APPLICATION NUMBER: 522,952
 FILING DATE: 03-APR-1990
 APPLICATION NUMBER: 416,306
 FILING DATE: 03-OCT-1989
 APPLICATION NUMBER: 412,816
 FILING DATE: 26-SEP-1989
 SEQ ID NO.: 2
 LENGTH: 783
 5422248-2

Query Match	7.1%;	Score 96;	DB 6;	Length 783;
Best Local Similarity	24.9%;	Pred. No. 0.043;		
Matches 54;	Conservative 34;	Mismatches 93;	Indels 36;	Gaps 12

```

OY      114 YRKTPTVFW -OMVNSFNAIVNYSNRS-----GDRPT--VRQGTAYVSATGAVATA 165
Db      562 Y-----TIFWTAQNQSFSAILINASSRGVLTGSLGEBASLYHILHMAASQAGATNSTVLLT 616
OY      166 LGKSLTTHLPVGRF---VPPAAVAANANCINIPLMQRELOVGIPIVADAGORIGSY 222
Db      617 MTLTPEGSELHTILGIFGLLLLTLCGTATWLCCSPNRKRPPLPSVP--DPAHSLSGSW 674
OY      223 -TAAQGIQOVVISRMCALPAMAPLIMLMDLEKDD 258
Db      675 PTIMEDDAFO-----LFGGTPTPTKLVLEED 702

```

RESULT 3
US-07-923-976-4
; Sequence 4, Application US/07923976
; Patent No. 5574136
; GENERAL INFORMATION:

APPLICANT: Nagata, Shigeazu
 APPLICANT: Fukuoka, Rikio
 TITLE OF INVENTION: DNA Encoding Granulocyte
 TITLE OF INVENTION: Colony-Stimulating Factor Receptor
 NUMBER OF SEQUENCES: 8
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Jones, Tullier & Cooper, P.C.
 STREET: P.O. Box 2266 Eads Station
 CITY: Arlington
 STATE: Virginia
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/923,976

Query Match	7.1%;	Score 96;	DB 1;	Length 836;
Best Local Similarity	24.9%;	Pred. No. 0.048;		
Matches 54;	Conservative 34;	Mismatches 93;	Indels 36;	Gaps 12;

```

QY      114 YRKPTVFW-OMVNSFNALVNYSNRS-----GDTPIF--VRDLGATVYSATGAAYATA 165
Db      562 Y-----TIPTWNAQNQSPSALINASSRGPVLHGLEPASYLHIHLMMAASQAGATNSVTLT 616
QY      166 LGKSLITKELPRVYGRF--VPPAAVAAANCINIPMYNQRELONGIPADADAGRLGYSV 222
Db      617 MTLTPGSGSEHIIILGIFGLILLTLCICGTAMVCSRNKKNPLMPSVP--DEAMSLSSWV 674
QY      223 -TAAQGIQGVYISRCMAIPMAPIPLIMDLBKDD 258
Db      675 PTMEEDARQ-----LPGLGTPPIITKLTVLEED 702

```

RESULT 4
US-07-923-976-8
; Sequence No. Application US/07923976
; Patent No. 5574136
; GENERAL INFORMATION:
; APPLICANT: Nagata, Shigekazu
; APPLICANT: Fukunaga, Rikio
; TITLE OF INVENTION: DNA Encoding Granulocyte

TITLE OF INVENTION: Colony-Stimulating Factor Receptor
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jones, Tullar & Cooper, P.C.
STREET: P.O. Box 2266 Eads Station
CITY: Arlington
STATE: Virginia
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/923,976
FILING DATE: 19920922
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 74539/1990
FILING DATE: 23-MAR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 176629/1990
FILING DATE: 03-JUL-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP91/00375
FILING DATE: 22-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Hellwege, James W.
REGISTRATION NUMBER: 28, 808
REFERENCE/DOCKET NUMBER: 514853
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-415-1500
TELEFAX: 703-415-1508
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 863 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-923-976-8

Query Match 6.7%; Score 89.5; DB 1; Length 863;
Best Local Similarity 22.4%; Pred. No. 0.29;
Matches 55; Conservative 35; Mismatches 90; Indels 65; Gaps 13;

QY 61 VTPGTEQDLRAK--YYDSAFHPTDGEKVL--IGRMSAQ--VPMNMTTTCMLTF 113
DB 503 IYTF-LYQDMGPQSHVAVYSGEMAPSHAPELHGHGKTAQLQEMVBRPPLGKPLTH 561
QY 114 YRKPTTVFV-OWVNSFNALVYNSNRSGDTPIT--VRQGTAVVSATTGAVALTA 165
DB 562 Y-----TFMTVAQNSFSALINASSRGFVLHGLEPASLYHILMAASQAGATNSTVLT 616
QY 166 LGLKSLTGLPLVGRFVFAVAANAACI-----NIPLMRQRELQVG 207
DB 617 MLTPRGSELHITLGL--FGILLITLCTGTAATCCSEPNRNPIMPSPDPDASHLSLGSW 673
QY 208 IPVADE-----AGRLGYSVTAAK-----OGIFQVVISRICMAIPMAIPPLIMDT 253
DB 674 VPTIMEELRPGRQGMGLQGTSEMSRALTPHPCVDQAFQ-----LPGIGTPITITLT 724
QY 254 LKKKD 258
DB 725 VLEED 729

RESULT 5
US-09-134-001C-5531
Sequence 5531, Application US/09134001C
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 5531
LENGTH: 655
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5531

Query Match 6.6%; Score 89; DB 4; Length 655;
Best Local Similarity 20.9%; Pred. No. 0.21;
Matches 54; Conservative 32; Mismatches 78; Indels 94; Gaps 10;

QY 53 IVQYRAGVTPGTEQDLRAKYYSDFHPTDGEKVLIGRMSAQVPMNMTTTCMLT 112
DB 407 IIAFLAGYLTOGI-----KYTRRLPQAIBGLKPTLIYPL-----LSVSIITGLMV 453
QY 113 FYRKPTTVFVQWVNSFNALVYNSNRSGDTPITVROL-----GTAYVSA 157
DB 454 YVFNPPAA---WLN---HLLNLGINSLSGSNTMLGLVIGAMMAIDMGFPNKAAYVFA 506
QY 158 TTG-----AVATA-----LGLKSLTKHL 175
DB 507 TLAITEGNAAPITAMIGMIPPLAIAFMILFRKFTKEGRGSIVPVYVNGLSITBGA 566
QY 176 PPLVGRFVFPAAVAANCINIPLMRQRELQVGPVADAGORLGYSVTAAGQIFQV-- 233
DB 567 -----IPFAAAD-----PLRVIPSMVSGVAGAIALGLSSIRAPHGIFVITGT 612
QY 234 -ISRICMAIPMAIPPLI 250
DB 613 DFNHILQTLIALVVGTLV 630

RESULT 6
US-09-252-991A-18502
Sequence 18502, Application US/09252991A
Patient No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 18502
LENGTH: 394
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18502

Query Match 6.5%; Score 88; DB 4; Length 394;
Best Local Similarity 18.7%; Pred. No. 0.12;
Matches 46; Conservative 40; Mismatches 64; Indels 96; Gaps 11;

QY 35 TDPRTLLSGAQLSASRNIVQYRAGVTPGTEQDLRAKYVDSA----- 81
DB 139 SEPLVQIVTDSYPTATYVENVYARGV-----QFVRAG--LDVAAPAAQVMLEPRYW 189
QY 82 FHPDTEKVLIGRMSAQVPMNMTTTCMLTFYRKPTTVFVQWVNSFNALVYNSNRSG 141
DB 190 FNPLESRRALI---PGALIVMTITIGTML-----TALVVARBWRGTMENVLS----- 235

ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/316,424A
FILING DATE: 30-SEP-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1331-122
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 146 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-316-424A-6

Query Match 6.1%; Score 82; DB 3; Length 146;
Best Local Similarity 25.8%; Pred. No. 0.13;
Matches 32; Conservative 19; Mismatches 57; Indels 16; Gaps 4;

QY 120 VVFWQVNSFNATVNSRSGDPTITVROLGTAVYSATGAVATAL--GKSLTKLHP 177
DB 32 LVVFWPTGRYFDSF-----GDLSSASATGNKVKHKKVITAFNDGL---NHLDS 80

QY 178 LVGRFVPPAAVAANCINIPIMRORELOVGIPVADAGORLGSYTAAGKQIFQVVISRI 237
DB 81 LKG---TFASLSLHCXKLVHPENFRLLGNMIVIVLGHHLGKQFTPAQAQAFQVAVAGV 137

QY 238 CMAI 241
DB 138 ATAL 141

RESULT 10
US-09-005-546-20
Sequence 20, Application US/09005546
Patent No. 6090782
GENERAL INFORMATION:
APPLICANT: TSYRLOVA, IRENA
APPLICANT: WOLPE, STEPHEN D.
TITLE OF INVENTION: INHIBITOR OF STEM CELL PROLIFERATION AND
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/005,546
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/535,882
FILING DATE: 28-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1331-177
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 146 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-005-546-20

Query Match 6.1%; Score 82; DB 3; Length 146;
Best Local Similarity 25.8%; Pred. No. 0.13;
Matches 32; Conservative 19; Mismatches 57; Indels 16; Gaps 4;

QY 120 VVFWQVNSFNATVNSRSGDPTITVROLGTAVYSATGAVATAL--GKSLTKLHP 177
DB 32 LVVFWPTGRYFDSF-----GDLSSASATGNKVKHKKVITAFNDGL---NHLDS 80

QY 178 LVGRFVPPAAVAANCINIPIMRORELOVGIPVADAGORLGSYTAAGKQIFQVVISRI 237
DB 81 LKG---TFASLSLHCXKLVHPENFRLLGNMIVIVLGHHLGKQFTPAQAQAFQVAVAGV 137

QY 238 CMAI 241
DB 138 ATAL 141

RESULT 11
US-08-477-669-6
Sequence 6, Application US/08477669
Patent No. 6432917
GENERAL INFORMATION:
APPLICANT: KOZLOV, VLADIMIR
APPLICANT: TSYRLOVA, IRENA
TITLE OF INVENTION: INHIBITOR OF STEM CELL PROLIFERATION AND
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,669
FILING DATE: 07-Jun-1995
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/316,424
FILING DATE: 30-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1331-122
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100

INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 146 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-08-477-669-6

Query Match
Best Local Similarity 25.8%; Score 82; DB 4; Length 146;
Matches 32; Conservative 19; Mismatches 57; Indels 16; Gaps 4;

Db 120 VTFVQWVQSFNAIVNYSNRSQDPTITVROLGTAVYSATTGAVATL-GIKSLTKLPP 177
32 LVVYPTQRYFDSF-----GDLSSASAIMGNKAKVAKHGKIVITAFNDGL---NHLDS 80
QY 178 LVGRFVFPAAVAANCINIPLMQRELOVGIPIVADGQRLGYSVTAKGQIFQVVISRI 237
81 LKG---TFASLSLHCDKLVHPDENFRLLGNMIVIVLGHGKQDTPAAQAAFOKVAVG 137
Db 238 CMAI 241
QY 138 ATAL 141

RESULT 12
US-07-923-976-6
Sequence 6, Application US/07923976
Patent No. 5574136
GENERAL INFORMATION:
APPLICANT: Nagata, Shigekazu
APPLICANT: Fukunaga, Rikito
TITLE OF INVENTION: DNA Encoding Granulocyte
TITLE OF INVENTION: Colony-Stimulating Factor Receptor
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jones, Tulliar & Cooper, P.C.
STREET: P.O. Box 2266 Bads Station
CITY: Arlington
STATE: Virginia
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/923,976
FILING DATE: 19920922
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 74539/1990
FILING DATE: 23-MAR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 176629/1990
FILING DATE: 03-JUL-1990
APPLICATION NUMBER: PCT/JP91/00375
PRIOR APPLICATION DATA:
FILING DATE: 22-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Helliwege, James W.
REGISTRATION NUMBER: 28,808
REFERENCE/DOCKET NUMBER: 514853
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-415-1500
TELEFAX: 703-415-1508
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 771 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear

MOLECULE TYPE: protein
US-07-923-976-6

Query Match
Best Local Similarity 29.1%; Score 82; DB 1; Length 771;
Matches 41; Conservative 21; Mismatches 57; Indels 22; Gaps 9;

QY 61 VTPGITEPOLMRK--YYDSAFHPDTGKVT--IGKSAQ---VPMNMITTCMLTF 113
DB 503 IYTP-LYODTMSQHVYAVSQEMAPSHAPLHKIKIGKQWQLEWVPEPELKGSPLE 561
QY 114 YRKTPTVTFM-QWVNSFNAIVNYSNRS---GDTPTVROLGTAVYSATTGAVATLGL 168
DB 562 Y-----TTFMTNAQNSFSLHNASRGVTLGEPASLVHILMAASQAGATNSTVTL 616
QY 169 KSLTKHLPPLVGRFVPAAVA 189
DB 617 MTLT---PAPYGR-IPSGQVS 633

RESULT 13
US-09-252-991A-25845
Sequence 25845, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 25845
LENGTH: 514
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25845

Query Match
Best Local Similarity 23.5%; Score 81; DB 4; Length 514;
Matches 35; Conservative 24; Mismatches 52; Indels 38; Gaps 6;

QY 118 PTVFQWVQSFNA-----IVNYSNRSQDPTITVROLGTAVYSATTGAVATLGIKSL 171
DB 376 PALAWFAPVPLVGCANNALLARAYAAQDTRPLRCELGSALNALAVLPLIFGLPG- 434
QY 172 TKHLPVGRFVPAAVAANCINIPLMQRELOVGIPIVADGQRLGYSVTAKGQIFQ 231
DB 435 -----IPLAALGVLTCTALLMRQALGLPVA-----RL-WLSALNALAA 477
QY 232 VVISRI-----CMAIPMAI 246
DB 478 GLIFRIDGIWLGGLGTAVAGCLLALGLAL 506

RESULT 14
US-08-947-965-71
Sequence 71, Application US/08947965A
Patent No. 6004790
GENERAL INFORMATION:
APPLICANT: Dijkhuizen, Lubbert
APPLICANT: Dijkstra, Bauke
APPLICANT: Andersen, Carsten
APPLICANT: Oester, Claus von der
TITLE OF INVENTION: Cyclomalto-dextrin Glucanotransferase
FILE REFERENCE: 4285.204-US
CURRENT APPLICATION NUMBER: US/08/947,965A
CURRENT FILING DATE: 1997-10-09

EARLIER APPLICATION NUMBER: 0477/95
EARLIER FILING DATE: 1995-04-21
EARLIER APPLICATION NUMBER: 1173/95
EARLIER FILING DATE: 1995-10-17
EARLIER APPLICATION NUMBER: 1281/95
EARLIER FILING DATE: 1995-11-16
EARLIER APPLICATION NUMBER: PCT/DK96/00179
EARLIER FILING DATE: 1996-04-22
NUMBER OF SEQ ID NOS: 78
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 71
LENGTH: 676
TYPE: PRT
ORGANISM: Bacillus sp.
US-08-947-965-71

Query Match 6.0%; Score 81; DB 3; Length 676;
Best Local Similarity 20.5%; Pred. No. 2;
Matches 43; Conservative 33; Mismatches 82; Indels 52; Gaps 9;

QY 67 TEDQIMRAKIVYDSAFHPTDGEKVLIGMSAQVPMNTIT-----107
DB 405 TERIMEDIYIYERTF---GNSIVLTAVNSN--SNQITITLNTSLPQGNVTDLQORL 458
QY 108 -----GCMLEFYRKPTPVFWQVWVNSFNAIVN-----YSNRSGDTPITVR-----148
DB 459 DGNITTVANGAVNSFQLRANSVAWQVSNPSTPLIGQVPMGKAGNT-ITVSGGFG 517
QY 149 -QLGTAIVSATTAATGATATLGLKSLTKHLPLVGRFVPPAAVAANCINIPLMRQRELQVG 207
DB 518 DERGSVLFDSSTSEISNNTKISVK-VPNVAGGYDLSVTAAN-IKSPYKEREVLG 575
QY 208 IPVADEAGORLGYSVTAAGIFQVVISRI 237
DB 576 ----NQVSVRFQVNNATTSPGTNLIVGNV 601

RESULT 15
US-09-252-991A-19749
Sequence 19749, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 19749
LENGTH: 621
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19749

Query Match 5.9%; Score 80; DB 4; Length 621;

Best Local Similarity 21.9%; Pred. No. 2.3;
Matches 39; Conservative 26; Mismatches 53; Indels 60; Gaps 9;

QY 88 EKVVLIGMSAQVPMNTITGCMLEFYRKPTPVFWQVWVNSFNAIVVNSNRSGDTPITV 147
DB 287 DAIVLIGRVLASAFSVMAVGFASLTYSRTAE---QWQDQ-----TPDSV 328
QY 148 ROLGTAIVSATTAATGATATLGLKSLTKHLPLVGRFVPPAAVA-AANCINIPLMRQRELQ- 205
DB 329 RY-----ALGMDARKLGPPSVN---LNAVAKASGTVDLPRLTNEARG 369
QY 206 -----VGIPVADEAGORLGYSVTAAGIFQVVISRICMAIPMAIPLIM 251

DB 370 NTTLSVSTDGVSVPKAVPV-RMAAYNATT---GLYEVTVPS-----TTAEAPPLIL 418

Search completed: February 3, 2004, 09:15:15
Job time : 23 secs

A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-324 <WIL>
 A/Cross-references: EMBL:Z66565; PIDN:CAA91477.1; GSPDB:GN00020; CESP:T04F8.1
 A:Experimental source: clone T04F8
 C/Genetics:
 A:Gene: CESP:T04F8.1
 A:Map position: X
 A:Intons: 29/3; 58/2; 172/3; 201/2; 244/1
 C/Superfamily: Saccharomyces probable membrane protein YOR271c

Query Match 55.2%; Score 742.5; DB 2; Length 324;
 Best Local Similarity 55.6%; Pred. No. 1.6e-57;
 Matches 145; Conservative 43; Mismatches 72; Indels 1; Gaps 1;

QY 1 MESKGEPLDINIOEPRMDSTFLGRARHFFVTDPNNLLSGAOLEASRNIVQNRAG 60
 Db 1 MSRLVKTIVLRPDISKRMQDSTFEGRAKFFAITNPLNLFHEKQDEKRLVEDIRK 60
 QY 61 VTPFGITDQLMRAKYVDSAFHPTGEXVLLIGMSAQVPMNMTTTCMLTFYRKTPTV 120
 Db 61 SVSNDLTLNQLMRAKYVDSAFHPTGEXVLLIGMSAQVPMNMTTTCMLTFYRKTPTV 120
 QY 121 VFMQWVNSFNALVYNSGSDPTITROGLTAVSTAGVATGALGSLTKLPLVNG 180
 Db 121 IFWQMLNQSFNAYVNSGSDG-SVQLVSTCAATGALTAALGSLVKKAPPLVG 179
 QY 181 RFPFPAVAANANCINIPLMRQELQVGIPIVADEAGRLGYSVTAKQGIPOVVISRICMA 240
 Db 180 RLVPFVAVCANSNIPMRGELTBIGDILDENGVIGSGPGVAGSAISQVVSRIEMA 239
 QY 241 IPAMAIPLIMDTLEKKDFLK 261
 Db 240 VPSFAFIPVVVNALEKRPYFK 260

RESULT 3

hypothetical protein C47D12.3 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
 C/Accession: T19996
 R/Gajadaty, S.
 submitted to the EMBL Data Library, March 1996
 A/Reference number: Z19209
 A/Accession: T19996
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-326 <WIL>
 A/Cross-references: EMBL:Z69902; PIDN:CAA93764.1; GSPDB:GN00020; CESP:C47D12.3
 A:Experimental source: clone C47D12
 C/Genetics:
 A:Gene: CESP:C47D12.3
 A:Map position: 2
 A:Intons: 26/1; 116/2; 205/2
 C/Superfamily: Saccharomyces probable membrane protein YOR271c

Query Match 51.5%; Score 693; DB 2; Length 326;
 Best Local Similarity 53.7%; Pred. No. 3.6e-53;
 Matches 139; Conservative 40; Mismatches 73; Indels 6; Gaps 2;

QY 3 SKNGELPLDINIOEPRMDSTFLGRARHFFVTDPNNLLSGAOLEASRNIVQNRAGV 62
 Db 6 SKCTELP---DISPRWDONTFOGRVAVYFSTANCLMLFVSNALTEKARIVLEVKQGY 62
 QY 63 TPGITDQLMRAKYVDSAFHPTGEXVLLIGMSAQVPMNMTTTCMLTFYRKTPTV 122
 Db 63 DPNMTVDELAKAKTLVYSARHPDGEKMFILGRISAQVPCMLITGMLTFYQCLPHYIF 122
 QY 123 WQWVNSFNALVYNSGSDPTITVROGLTAVSTAGVATGALGSLT---KHLPLV 179
 Db 123 FHWVNSFNALVYNSGSDPTITVROGLTAVSTAGVATGALGSLT---KHLPLV 182

QY 180 GRFVPPAAVAANANCINIPLMRQELQVGIPIVADEAGRLGYSVTAKQGIPOVVISRICM 239
 Db 183 ARLVPPAAIAFANAIINIPMRNKEFTNGIPVEDBGRITGFSYVAGHALPOVVISRVGM 242
 QY 240 AIPAMAIPLIMDTLEK 256
 Db 243 AVNNVVLGPIVLEQLSK 259

RESULT 4

hypothetical protein AH6.2 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
 C/Accession: T18612
 R/Berke, M.
 submitted to the EMBL Data Library, January 1995

A/Reference number: Z18998
 A/Accession: T18612
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-329 <WIL>
 A/Cross-references: EMBL:Z48009; PIDN:CAA88076.1; GSPDB:GN00020; CESP:AH6.2
 A:Experimental source: clone AH6
 C/Genetics:
 A:Gene: CESP:AH6.2
 A:Map position: 2
 A:Intons: 26/1; 99/3; 173/3; 202/2; 245/1
 C/Superfamily: Saccharomyces probable membrane protein YOR271c

Query Match 47.5%; Score 639; DB 2; Length 329;
 Best Local Similarity 49.0%; Pred. No. 2e-48;
 Matches 122; Conservative 44; Mismatches 83; Indels 0; Gaps 0;

QY 13 NIOEPRMDSTFLGRARHFFVTDPNNLLSGAOLEASRNIVQNRAGVTPGITEDQLM 72
 Db 13 DISKPKDQTYGRARHFFSTINPLTFSSRQOEKCEIVNNYTGVISPLTIDELW 72
 QY 73 RAKYVDSAFHPTGEXVLLIGMSAQVPMNMTTTCMLTFYRKTPTVFMQWVNSFNA 132
 Db 73 KAKTLVDSYHPDTEKMFILGRISAQVPCMLITGMLTFYQCLPHYIF 132
 QY 133 IVNYSRSGDPTITVROGLTAVSTAGVATGALGSLTKLPLVGRFVPPAAVAAN 192
 Db 133 VVNYTRSGNSKATNERLTFVSTCATSGAFTVGLGNKVKNSHGLARLVPPAAIALAN 192
 QY 193 CINIPLMRQELQVGIPIVADEAGRLGYSVTAKQGIPOVVISRICMAIPAMAIPLIMD 252
 Db 193 AINIPMRNSNEASEGHELDENDQLVGKSKMAALSIAGVTLSRIMAMPYVMTPIIN 252
 QY 253 TLEKKDFLK 261
 Db 253 RITRTAYYR 261

RESULT 5

hypothetical protein F37H8.4 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
 C/Accession: T21924
 R/Gregory, J.
 submitted to the EMBL Data Library, November 1996

A/Reference number: Z19488
 A/Accession: T21924
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-326 <WIL>
 A/Cross-references: EMBL:Z81534; PIDN:CAB04347.1; GSPDB:GN00020; CESP:F37H8.4
 A:Experimental source: clone F37H8
 C/Genetics:
 A:Gene: CESP:F37H8.4
 A:Map position: 2

A:Introns: 26/1; 99/3; 171/3; 200/2; 243/1; 298/3
 C:Superfamily: Saccharomyces probable membrane protein YOR271c

Query Match 44.9%; Score 604; DB 2; Length 326;
 Best Local Similarity 47.3%; Pred. No. 2.3e-45;

Matches 121; Conservative 47; Mismatches 86; Indels 2; Gaps 1;

QY 1 MESKQGELELDINIDEPKRDSTFLGRARHFTVDPRLNLLSGAOLEASRNIVQYRAGVTPGITEEDOLM 60
 DB 1 MSELVSNLKTDPDISKPOQSRYYGRVHFPTLTNPILTLTSSVAKQECROIADLYKNG 60
 QY 61 VTPGITEEDOLMRAKYVDSAFHPDTEKVLIGRMSAOVPMNMITTCMLTFYKTPV 120
 DB 61 KVSPLITVSELMKATLVDSTYHPDTEKVLIGRMSAOVPMNMILNGLSLVYTFPV 120
 QY 121 VFWQVNOSEFNAIVYNSGSDPTTTRQLGTAVYSATTGAVALGKSLTKHLPLVVG 180
 DB 121 VFSHINOSFNAVYVYNSGNSKTSNERLILSYCATGAAALSLNANKN--SIA 178
 QY 181 RFPVPAVAANNCINIPLMRORELQVGPVADAGQRLGYSTAAKQGFQVVISRICMA 240
 DB 179 RLVPFAAVALANTINIPMIRSNVEGTELRDENGELLARSHQMAILSLAQVTLSRIMA 238
 QY 241 IPMAIPLIMDTLEK 256
 DB 239 MPDMVTPIIMRITR 254

RESULT 6

727337

hypothetical protein Y6E2A.9 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000

C:Accession: T27337

A:Submitted to the EMBL Data Library, January 1998

A:Reference number: Z20347

A:Accession: T27337

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-329 <MIL>

A:Cross-references: EMBL:AL021175; PIDN:CAA15970.1; GSPDB:GN00023; CESP:Y6E2A.9

A:Experimental source: clone Y6E2A

C:Genetics:

A:Gene: CESP:Y6E2A.9

A:Map position: 5

A:Introns: 26/1; 99/3; 202/2; 245/1

C:Superfamily: Saccharomyces probable membrane protein YOR271c

Query Match 44.1%; Score 593; DB 2; Length 329;
 Best Local Similarity 46.6%; Pred. No. 2.2e-44;

Matches 116; Conservative 39; Mismatches 94; Indels 0; Gaps 0;

QY 13 NIOEGRMOSQTFLGARHFTVDPRLNLLSGAOLEASRNIVQYRAGVTPGITEEDOLM 72
 DB 13 DISKRMIDLITSGKAYHIFASANPMTLFTSSNTQEMCKIIVDYKGIINDELIMDEL 72
 QY 73 RAKYVYDSAFHDPDEKVLIGRMSAOVPMNMITTCMLTFYKTPVTFWQVNOSEFNA 132
 DB 73 SAKIIVDSVYHPDTEKVLIGRMSAOVPMNMITTCMLTFYKTPVTFWQVNOSEFNA 132
 QY 133 IYVYNSGSDPTTTRQLGTAVYSATTGAVALGKSLTKHLPLVVGFPVPAVAAN 192
 DB 133 IYVYNSGSDPTTTRQLGTAVYSATTGAVALGKSLTKHLPLVVGFPVPAVAAN 192
 QY 193 CINIPLMRORELQVGPVADAGQRLGYSTAAKQGFQVVISRICMAIPMAIPLIMD 252
 DB 193 AINIPMIRSNVEGTELRDENGELLARSHQMAILSLAQVTLSRIMAMPDMVTLSPVIMN 252
 QY 253 TLEKQDFLK 261
 DB 253 RFTRTAYYK 261

RESULT 7

T15498

hypothetical protein C14F5.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 18-Feb-2000

C:Accession: T15498

R:Mix, P. submitted to the EMBL Data Library, June 1995

A:Description: The sequence of C. elegans cosmid C14F5.

A:Reference number: Z18361

A:Accession: T15498

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-285 <MIN>

A:Cross-references: EMBL:U29082; NID:g861384; PID:g861388; PIDN:AAA6404.1; CESP:C14F5.4

A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:C14F5.4

A:Introns: 31/3; 82/2; 136/2; 227/2

C:Superfamily: hypothetical protein C14F5.4

Query Match 35.9%; Score 483.5; DB 2; Length 285;
 Best Local Similarity 44.4%; Pred. No. 7.3e-35;

Matches 103; Conservative 37; Mismatches 89; Indels 3; Gaps 2;

QY 16 EPRWDQSTFLGRARHFTVDPRLNLLSGAOLEASRNIVQYRAGVTPGITEEDOLMRAK 75
 DB 40 QPRYDQSTFYGRRLRFAGMTDPLAFSSSTTELITASBLQCKRKKPVPA--TLBELHRSQ 98
 QY 76 YVYSAFHPDDEKVLIGRMSAOVPMNMITTCMLTFYKTPVTFWQVNOSEFNAIVN 135
 DB 99 RLQSAFHPDDEKVLIGRMSAOVPMNMITTCMLTFYKTPVTFWQVNOSEFNAIVN 158
 QY 136 YSNRSGDPTTTRQLGTAVYSATTGAVALGKSLTKHLPLVVGFPVPAVAAN 193
 DB 159 YTNRAKSTLTITTKOLVYSYRVSAGALMAIGKTYFPAKQSSPLAQRLVPLGAVAVANA 218
 QY 194 INIPLMRORELQVGPVADAGQRLGYSTAAKQGFQVVISRICMAIPMAI 245
 DB 219 INIPMIRSNVEGTELRDENGELLARSHQMAILSLAQVTLSRIMAMPDMVTLSPVIMN 270

RESULT 8

S67168

probable membrane protein YOR271c - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein O5433

C:Species: Saccharomyces cerevisiae

C:Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 19-Apr-2002

C:Accession: S67168; S67173; S72042

R:Jauliaux, J.C.; Polrey, R.

A:Submitted to the Protein Sequence Database, July 1996

A:Reference number: S67143

A:Accession: S67168

A:Molecule type: DNA

A:Residues: 1-327 <JAU>

A:Cross-references: EMBL:Z75179; NID:g1420607; PID:g1420608; MIPS:YOR271c

A:Experimental source: strain S288C

R:Cheret, G.; Sor, F.

A:Submitted to the Protein Sequence Database, July 1996

A:Reference number: S67169

A:Accession: S67173

A:Molecule type: DNA

A:Residues: 1-327 <CHB>

A:Cross-references: EMBL:Z75179; NID:g1420607; PID:g1420608; MIPS:YOR271c

R:Cheret, G.; Bernard, A.; Sor, F.

A:Yeast 12, 1059-1064, 1996

A:Title: DNA sequence analysis of the VP1-SNP2 region on chromosome XV of Saccharomyces

A:Reference number: S72039; MUID:97051594; PMID:8896271

A:Accession: S72042

A:Status: nucleic acid sequence not shown; translation not shown

A:Residues: 1-327 <CHW>
A:Cross-references: EMBL:X89633, NID:g1279694, PIDN:CAA61777.1, PID:g1279698
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1995
C:Genetics:
A:Cross-references: SGD:S0005797
A>Note: position: 15R
A>Note: YOR271c
C:Superfamily: Saccharomyces probable membrane protein YOR271c
C:Keywords: transmembrane protein
E:184-200/Domain: transmembrane #status predicted <TM1>
E:276-292/Domain: transmembrane #status predicted <TM2>

	Query Match	34.5%	Score 464.5:	DB 2,	Length 327;
	Best Local Similarity	40.6%;	Pred. No. 4e-33;		
	Matches 104; Conservative	41;	Mismatches 102;	Indels 9;	Gaps 2
Oy	8 LPIDINIEPMDOSTFLGRARHFFVTVDPRMLLSGAOLEASRNIVQNYRAGVTPGAT	67			
	: :				
Dd	5 VPGEIDLESKYDSTWGRIRHCNEISPTMLLTTEKDLAKAREIISAIRYGELEKE--T	62			
	: :				
Oy	68 EDOLMRKKYVDSAEHPDTGEKRVILGRNSAOVPMMNTITTCMLTFYRKPTFVFWQWN	127			
	: :				
Dd	63 TPEFRARKQKDSTVTHPDTGKTVLLPFRRSSNVLSLVTYGMLTPTGGCTAGTVFMQWAN	122			
	: :				
Oy	128 QSFNAIVNYSNRSGDTPITTRQLGTAYIASATTCAVAATALG-----LKSLTKHLPELVG	180			
	: :				
Dd	123 QSLNAVVASANANKSHPMSTQSOLLTNVAAAVTASCGVALGLNNLVRLENINISPHSRTLLG	182			
	: :				
Oy	181 RFVPRAAVAANNCIIPLIMRORELQVGPVADAEQRGLGYSTPAKKGIFQVVISRICMA	240			
	: :				
Dd	183 RLVPRAAIVASGIIVAVFLMRGBEIRKIGSVSPDSNGDEGVSKSAAPMAVGELASRLVINA	242			
	: :				
Oy	241 IPAMAIPLIMDTLEK	256			
	: :				
Dd	243 TPTMTVIPPLILVRLQR	258			

RESULT 9
T19873
hypothetical protein C41C4.2 - *Caenorhabditis elegans*
C|Species: *Caenorhabditis elegans*
C|Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C|Accession: T19873
R|Burton, J.
submitted to the EMBL Data Library, February 1995
A|Reference number: Z19190
A|Accession: T19873
A|Status: preliminary; translated from GB/EMBL/DBDJ
A|Molecule type: DNA
A|Residues: 1-616 <MWL>
A|Cross-references: EMBL:Z48045; PIDN:CAA88099.1; GSPDB:GN00020; CESP:C41C4.2
A|Experimental source: clone C41C4
C|Genetics:
A|Gene: CESP:C41C4.2
A|Map position: 2
A|Intons: 24/2; 72/3; 214/3; 291/2; 371/2; 419/1; 481/3; 503/1; 530/2

Query Match	32.2%	Score 432.5	DB 2	Length 616
Best Local Similarity	40.0%	Pred. NC.5.7e-30		
Matches 100	Conservative 43	Mismatches 96	Indels 11	Gaps 6
Qy	16	EPRMDSQTFILGRARHFTVTDPFRNLLEGAQQLSEASRNITQYRAVVTTCGITEDQIMPAK	75	
Db	44	EPRPQDFLGRKYLHCLVIDPRITLFAFNKCLBESLEJLNSKATATLNVEDSKSLMBAQ	102	
Qy	76	VYVDSAFHPDGEKVLIGRNSAQVPMNMTTGMCLTFYRKTPTVFMQVYNQSFNAIV	134	
Db	103	KLKSAIILHPDGEKVLPEFRMSGFVPFGWITVTGMLPNSPMPITLFMQWNQSHNACV	161	
Qy	135	NYSNNSGDTPIITVRDLGTAYVASATGAVATLGLSLTK---HLDP---LVGRVPPAA	187	
Db	162	NYANNNATQPOPULSKYIGAYGAALVPAACISGGLTYFTFKASSLPPTTRIIILQRFVPLPA	221	

Qy	188	VAANACINIPILKROBELQVIGIPVAD-EAGOSRLGYSVTAKOGIFQVVISRCMAIPMAI	246
Db	222	TSLASSLANVICRMNELGTGVEYKDGKVVGVSVKVAAGVADTMMVRAFLVPIILLM	281
Qy	247	PRLIMDTLEK	256
Db	282	PFCIMPYLER	291

RESULT 10
T37847
probable transporter - fission yeast (*Schizosaccharomyces pombe*)
C|Species: *Schizosaccharomyces pombe*
C|Date: 03-Dec-1999 #sequence_rev1sion 03-Dec-1999 #text_change 21-Jan-2000
C|Accession: T37847
R|Murphy, U.; Harris, D.; Barrel, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A|Reference number: Z21749
A|Accession: T37847
A|Status: preliminary; translated from GB/EMBL/DBJ
A|Molecule type: DNA
A|Residues: 1-325 <NR>
A|Cross-references: EMBL:Z29162; PIDN:CAP16226.1; GSPDB:GN00066; SPDB:SPAC1766.15C
A|Experimental source: strain 972h-; contid c17G6
C|Genetics:
A|Gene: SPDB:SPAC1766.15C
A|Map position: 1
A|Intons: 53/2, 70/1, 94/3, 206/3
C|Superfamily: Saccharomyces probable membrane protein YOR271C

Query Match	30.6%	Score	411.5	DB	2	Length	325
Best Local Similarity	36.3%	Pred. No.	1.8e-28				
Matches	89	Conservative	48	Mismatches	99	Indels	9
						Gaps	2
QY	24	FLGRARHFFTTVDPRNLISGAQLAESKNIVONTBAGVYVTCITIEDQLMRKAYVDSAFH	83				
	:	:	:	:	:	:	:
	:	:	:	:	:	:	:
Db	19	FMLMRHADINDPRLLTSTQDLSNAKVTIEDYGAKRIAC--DDEVYMAKKIVDSTLH	76				
QY	84	PDTEGEKVLVIGRMSAQVPANMTTTCGMLTFPRKTFPVVFMQVNASFPAIVNYSNRSGDT	143				
	:	:	:	:	:	:	:
	:	:	:	:	:	:	:
Db	77	PDTEKEVFLFPRNSCVLNLNLVVTAGMLQPNLIGTAGIVFWMQMQSVYVAFAVNSNANKST	136				
	:	:	:	:	:	:	:
	:	:	:	:	:	:	:
QY	144	PITVRQLGTAIVSATTGAVATALGLKSLTKHL-----PPLVGRFVFPFAVVAANAINI	156				
	:	:	:	:	:	:	:
	:	:	:	:	:	:	:
Db	137	QLTLPMQTSXYVAVASASCGVAIGLTKIVPRNNPLSSSSKAVLRLTPFAAVASAGLVN	196				
	:	:	:	:	:	:	:
	:	:	:	:	:	:	:
QY	197	PLNRORELQVGPVADENAGQRIGSVTAAKOGIFQVVISRICALIIPMAIPLIMDTLEK	256				
	:	:	:	:	:	:	:
	:	:	:	:	:	:	:
Db	197	FLNRGSELNQIDIVDFKGEBSLGRSKKAAFAVAGETALSRIYNSPIWIVPVLVLMRLQK	256				
	:	:	:	:	:	:	:
	:	:	:	:	:	:	:
QY	257	KDPLK	261				
	:	:	:	:	:	:	:
	:	:	:	:	:	:	:
Db	257	QNNLR	261				

```

RESULT 11
T02841
mitochondrial tricarboxylate carrier MTCC [imported] - Leishmania major (strain Friedlin
C:Species: Leishmania major
C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 19-May-2000
C:Accession: B81461; T02841
R:Myer, P. J.; Audleman, L. J.; deVos, T.; Hixson, G.; Kiser, P.; Lemley, C.;
Proc. Natl. Acad. Sci. U.S.A. 96, 2902-2906, 1999
A>Title: Leishmania major Friedlin chromosome 1 has an unusual distribution of protein-c
A:Reference number: A81455; MUID:99178987; PMID:10077609
A:Accession: B81461
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-326 <Pyl>
A:Cross-references: GB:AE001274; NID:G3264850; PIDN:AMC24664.1; PID:ig1617564; GSPDB:GN00
A:Experimental source: strain MHOM/IL/81/Friedlin
C:Genetics:
A:Gene: MTCC

```

A:Map position: 1
C:Superfamily: Saccharomyces probable membrane protein YOR271c
C:Keywords: mitochondrion

Query Match	20.7%	Score 278;	DB 2;	Length 326;
Best Local Similarity	30.9%	Pred. No. 9, 2e-17;		
Matches	80;	Conservative	45;	Mismatches 106;
				Indels 28;
				Gaps 8

[illegible]

RESULT 12

A99514
 hypochlorite protein MYPU 0170 [imported] - Mycoplasma pulmonis (strain UAB CT1P)
 C:Species: Mycoplasma pulmonis
 C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001
 C:Accession: A99514
 R:Chambud, I.; Heiligg, R.; Ferris, S.; Barbe, V.; Samson, D.; Gallison, F.; Mosser, I.;
 Nucleic Acids Res. 29, 2145-2153, 2001
 A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulmonis
 A:Reference number: A99512; MUID:21267165; PMID:11353084
 A:Accession: A99514
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-791 <KIR>
 A:Cross-references: GB:AL445566; PID:G14089430; PID:GAC13190.1; GSPDB:GN00153
 A:Experimental source: strain UAB CT1P
 C:Genetics:
 A:Gene: MYPU 0170
 A:Genetic code: SGC3
 C:Superfamily: phosphotransferase system N-acetylglucosamine-specific enzyme II, phospho-
 lyzyme II, factor III homology

```

Query March 7.5%; Score 101; DB 2; Length 791;
Best Local Similarity 22.3%; Pred. No. 0.94;
Matches 59; Conservative 39; Mismatches 81; Indels 86; Gaps 15;

QY 36 DPNRLLSGAQL--EASRNIYONTRAGVYTPGITEDQLMRKAYVDS--AFHPD-- 85
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 146 DPESLIRLVGSSVGFPTLQTSVFGGISVGLT-----SYLVNKKFHKIGFHPAFS 194
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 86 --TGEKVVLRMSAQVPMNMNTITGCMILTFRKPTPTVFQWQVNS----- 129
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 195 FFAKRRFVPLVAIIAMPVLSLTF-----LLEFWVVKKGISVFETALGKVPYG 241
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 130 -----FNAIIVNYSNRSQDPTIV---RQLGTAIVGAT--TGAVA 163
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 242 IESPTFGIERSLTPGLHNVFYAPLWYSNAGDATALSGMQAGNPFVSAGPTANIT 301
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 164 TALGLKSTIKPLPLVGRFVPPAAVAAANCINIPLMRQ---ELQVGIPIVADEAGORIGY 220
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 302 QEL-IDSTITKADKVVGDSTGMQAVNSLN-FNVVSPFRQGSSEVQT-LRVLDFAQBERGI 358
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 221 SVTAAKGQIFQVIVISRICMAIPMA 245
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

Db 359 KLGRFLQGY-VFMS--MGLPAA 379

RESULT 13

УН0329

granulocyte colony-stimulating factor receptor D7 precursor - human
G/Species: Homo sapiens (man)
C/Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 05-Nov-1999
C/Accession: JH0329; S21608
R/Blasters: A.; Davis, T.; Curtis, B.M.; Gimpel, S.; Slims, J.E.; Cosman, D.; Park, L.; Sorrentino, M.P.
J./Exp. Med. 172, 1559-1570, 1990
A>Title: Expression cloning of a human granulocyte colony-stimulating factor receptor: a cDNA encoding a type III membrane protein
A/Reference number: JH0329; MIM:31079757; PMID:2147944
A/Accession: JH0329
A/Molecule type: mRNA
A/Residues: 1-783 <LNR>
A/Cross-references: GB:X55720; NID:G31698; PIDN:CMA39252.1; PID:G31699
A/Experimental source: Placenta
C/Keywords: glycoprotein; transmembrane protein
F1-24/Domains: signal sequence #status predicted <Sig->
F125-783/Product: granulocyte colony-stimulating factor receptor D7 #status predicted <MIM:>
F125-627/Domains: extracellular #status predicted <Extr->
F1628-563/Domains: transmembrane #status predicted <TMB->
F1634-783/Domains: intracellular #status predicted <INT->
F193-128, 134, 389, 474, 579, 610/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match	7.1%;	Score 96;	DB 2;	Length 783;
Best Local Similarity	24.9%;	Pred. No. 2.5;		
Matches 54;	Conservative 34;	Mismatches 93;	Gaps 12;	

```

Qy      114 YRKPTPVFW-QMVNQSFNAIVNYSNRS-----GDTPIF--VRQGLAYVSATTGAAYATA 165S
Db      562 Y-----TITWTAHQNSPSALINMSRGFVLHGLEPABLYHHIIMAAQSQAGATNSVTLL 616S
Qy      166 LGLKSLITKLPLPVGRF--VFPAAVAAANACINIPMRQRLQVGIPIVADAGRGLGVS 222S
Db      617 MTLTPGSELIHTILGIFGLLLLTLCGTAWICCSPNRKPLMPSPV--DPAHSLGSMV 674S
Qy      223 -TAKQGIQVOWISRIQMAIPMAIPMLINDTLEKQD 258
Db      675 PTIMEDAFQ-----LPGLTGPITKLTIVLEED 702

```

RESULT 14

T40

fructosyl amine - fission yeast (*Schizosaccharomyces pombe*)

C|Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C|Accession: T40295
R|Wood, V.; Rajandream, M.A.; Barrell, B.G.; Gilbert, H.; Duesterhoeft, A.
submitted to the EMBL Data Library, March 1998
A|Reference number: Z21918
A|Accession: T40295
A|Molecule type: DNA
A|Status: preliminary; translated from GB/EMBL/DBJ
A|Molecule type: DNA
A|Residues: 1..412 <MOO>
A|Cross-references: EMBL:AL02071; PIDN:CAAI7815.1; GSPDB:GN00067; SPDB:SPBC354.15
A|Experimental source: strain 972h-; contid c354
C|Genetics:
A|Gene: SPDB:SPBC354.15
A|Map position: 2

```

Query Match 7.0%; Score 94.5; DB 2; length 412;
Best Local Similarity 20.6%; Pred. No. 1.5;
Matches 42; Conservative 37; Mismatches 76; Indels 49; Gaps 8;

Qy 41 LLSGAQLASRNLYONTKAG-VTPGTTEDQLMPAKYVYDSAFPDPDGKAVLLGR---- 95
      :::::
Db 41 VIDSSVDNARIIRSDYADAVYCSMGIDALEEMWNTPLPEQGF--GSGLMFGGRDNVE 97

```

QY 96 ---NSAQVPMNMTITGCLTFYRKPTV--VFQWV-----NOSF 130
Db 98 YRDMSLR---NLTKMGVSAKFTTEIRKLPKMGELNDGAGVAFSSGMANAEQSV 154
QY 131 NAIIVNENRSGDPTITVROLGTAIVSATGAVATAGLKSITKHLPLVGRFV---FAA 187
Db 155 KSVVNY-----LAHAGVSFISGEGVEEITEENVKGVRTTGAYMAEKLIFA 204
QY 188 VAAANCINIPLMROBELQVGIPIVA 211
Db 205 TGAWTASLLPNDHTFRFLATGQPIVA 228

RESULT 15

I56551
neurotiximin - rat:
C/Species: Rattus norvegicus (Norway rat)
C/Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 19-May-2000
C/Accession: I56551
R/Struyk, A.F.; Canoll, P.D.; Wolfang, M.J.; Rosen, C.L.; D'Eustachio, P.; Salzer, J.L.
J. Neurosci. 15, 2141-2156, 1995
A/Title: Cloning of neurotiximin defines a new subfamily of differentially expressed neu
A/Reference number: I56551; MUID:95198094; PMID:7891157
A/Accession: I56551
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-344 <RES>
A/Cross-references: EMBL:U16845; NID:G755184; PIDN:AAA67445.1; PID:G755185
C/Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-termin

Query Match 6.7%; Score 89.5; DB 2; Length 344;

Best local similarity 24.8%; Pred. No.3.4; 95; Indels 35; Gaps 10;
Matches 52; Conservative 28; Mismatches 95;

QY 2 ESKMGEPLDINIOEPFW-DOSTFGARHFTVTPRNLLSGAQLEASRNI--VQNYR 58
Db 51 ESATLRCTIDNRVAVANLNSTILYAGNDKCL-DPRVLLSNTQYISIEIQNVDD 109
QY 59 AGVVTPTGITEQLMRAKTVYDSAFHPTGEXVVLIGMSAOV-----PNNMTIT 107
Db 110 EGPTCSVQTDN-----HPKT-SRVHLIVQVSPRIEISDISINEGNNISLT 156
QY 108 GCMLEFYRKPTVVFQWVNSFNAINVNSRSGDPTITVROLGTAIVSATGAVATAG 167
Db 157 -CIAT-GHPEPTVT-WRHISPKAVGFVSEDEYLEIGITRQSGEYECASNDVAAPYR 213
QY 168 LKSLTKHLPLV---GRFVPPAAVAANAC 193
Db 214 RVNVTVNYPPYISEAKGTGVGQKGTLC 243

Search completed: February 3, 2004, 09:13:48
Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2004, 09:11:53 ; Search time 17 Seconds
(without alignments)
721.998 Million cell updates/sec

Title: us-09-990-415a-2

Perfect score: 1345
Sequence: 1 MESMGELPLDINIQEPRWD.....PMAIPLIMDTLEKKDFLK 261

Scoring table: BL0SUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1326	98.6	321	1	SFX3_HUMAN
2	1272	94.6	321	1	SFX3_RAT
3	1269	94.3	321	1	SFX3_MOUSE
4	1089	81.0	322	1	SFX1_HUMAN
5	1064	79.1	322	1	SFX1_RAT
6	1059	78.7	322	1	SFX1_MOUSE
7	753	56.0	322	1	SFX2_MOUSE
8	749	55.7	322	1	SFX2_HUMAN
9	639	47.5	329	1	VP22_CAEEL
10	484.5	36.0	340	1	SFX5_HUMAN
11	484.5	36.0	342	1	SFX5_MOUSE
12	432.5	32.2	600	1	SNE2_CAEEL
13	113	8.4	313	1	SFX2_MOUSE
14	96	7.1	836	1	GCSR_HUMAN
15	89.5	6.7	344	1	NTRI_RAT
16	88.5	6.6	344	1	NTRI_HUMAN
17	88.5	6.6	344	1	NTRI_MOUSE
18	88	6.5	594	1	RCO3_NEUR
19	87.5	6.5	234	1	YE99_ARCFU
20	85	6.3	914	1	PERT_RAT
21	85	6.3	926	1	PERT_PIG
22	84	6.2	272	1	EUTC_PSEK
23	82.5	6.1	274	1	PTRD_KLEPN
24	82	6.1	146	1	HBI_MOUSE
25	82	6.1	836	1	VG26_BPMLS
26	81.5	6.0	531	1	UD16_MOUSE
27	81	6.0	519	1	NIFL_AZOVI
28	81	6.0	703	1	CDGT_BACS2
29	81	6.0	914	1	PERT_MOUSE
30	81	6.0	2505	1	FAS_RAT
31	80.5	6.0	387	1	MURG_ZYMO
32	80.5	6.0	393	1	XYIH_ECOLI
33	80.5	6.0	913	1	DEOL_CHYV2

34	80	5.9	501	1	DLDH_PEA
35	80	5.9	617	1	PYS1_PSEAE
36	80	5.9	688	1	PYS2_PSEAE
37	80	5.9	815	1	KINH_CAEEL
38	79.5	5.9	353	1	CEPU_CHICK
39	79.5	5.9	529	1	UD16_RAT
40	79.5	5.9	693	1	GUAA_HUMAN
41	79	5.9	337	1	G55A_CHICK
42	79	5.9	455	1	YVAD_BACSU
43	78.5	5.8	412	1	YVAD_ECOLI
44	78.5	5.8	470	1	YJIR_ECOLI
45	78	5.8	4349	1	PAT2_HUMAN

ALIGNMENTS

RESULT 1
SFX3_HUMAN STANDARD, PRT, 321 AA.
AC Q9EWM7; Q9NTP4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Sideroflexin 3.
GN SFXN3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Heath P.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RX MEDLINE=22388257; PubMed=12477932;
RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Schenker C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ueda T.B., Tohiyuki S., Abramo J.P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Wotley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywicki M.I., Skalska J., Small D.E.,
RA Scheraga A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SUBCELLULAR LOCATION.
RA Rabliloud T.;
RL Unpublished observations (FEB-2002).
CC -!- FUNCTION: Potential iron transporter.
CC -!- SUBCELLULAR LOCATION: Mitochondrial.
CC -!- SIMILARITY: Belongs to the sideroflexin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
CC or send an email to license@ebi.ac.uk).


```

CC -----
DR EMBL; AL133215; -, NOT ANNOTATED CDS.
DR EMBL; BC00124; AA00124.1; -.
DR InterPro; IPR004686; Mtc.
DR Pfam; PF03820; Mtc; 1.
DR ProDom; PD006986; Mtc; 1.
DR TIGRfam; TIGR00798; mtc; 1.
KM Transport; Iron transport; Iron; Mitochondrion; Transmembrane.
FT TRANSMEM 146 164 POTENTIAL.
FT TRANSMEM 174 194 POTENTIAL.
FT TRANSMEM 226 246 POTENTIAL.
FT TRANSMEM 266 286 POTENTIAL.
FT COMPACT 112 112 K -> Q (IN REF. 1).
SQ SEQUENCE 321 AA; 35503 MW; 54718C600B3D34BD CRC64;

Query Match 98.6%; Score 1326; DB 1; Length 321;
Best Local Similarity 100.0%; Pred. No. 4.5e-110;
Matches 257; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 MGELPLDINIOEPRWDSTFLGRARHFFVTYDPRNLLSGAQLASRNIVQNTAGVTP 64
DB 1 MGELPLDINIOEPRWDSTFLGRARHFFVTYDPRNLLSGAQLASRNIVQNTAGVTP 60
QY 65 GLEDDQMRKAYVYDSAFHPDGTGEKVLIGRMSAQPMMNTITGMLTFYRKTPTVVFQ 124
DB 61 GLEDDQMRKAYVYDSAFHPDGTGEKVLIGRMSAQPMMNTITGMLTFYRKTPTVVFQ 120
QY 125 WVNQSFNAIVNYSRSGDPTITVQGLTAVSATGAVATAGLKSITKGLPPLVGRFVP 184
DB 121 WVNQSFNAIVNYSRSGDPTITVQGLTAVSATGAVATAGLKSITKGLPPLVGRFVP 180
QY 185 PAAVAAANCINIPLMRQRELOVGIPTVADAGORLGISVTAAKQIGQVVISRICMAIPAM 244
DB 181 PAAVAAANCINIPLMRQRELOVGIPTVADAGORLGISVTAAKQIGQVVISRICMAIPAM 240
QY 245 AIPPLIMDTLEKKDFLK 261
DB 241 AIPPLIMDTLEKKDFLK 257

RESULT 2
SF3X_RAT STANDARD; PRT; 321 AA.
AC Q9UR72;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Sideroflexin 3.
GN SF3X3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA Maehama H., Kojima I.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Potential iron transporter.
CC -1- SUBCELLULAR LOCATION: Mitochondrial (by similarity).
CC -1- SIMILARITY: Belongs to the sideroflexin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL; AF276997; AAF78249.1; -.
CC InterPro; IPR004686; Mtc.
CC Pfam; PF03820; Mtc; 1.

```

```

DR ProDom; PD006986; Mtc; 1.
DR TIGRfam; TIGR00798; mtc; 1.
KM Transport; Iron transport; Iron; Mitochondrion; Transmembrane.
FT TRANSMEM 146 164 POTENTIAL.
FT TRANSMEM 174 194 POTENTIAL.
FT TRANSMEM 226 246 POTENTIAL.
FT TRANSMEM 266 286 POTENTIAL.
SQ SEQUENCE 321 AA; 35433 MW; 61D9D679A7E6124 CRC64;

Query Match 94.6%; Score 1272; DB 1; Length 321;
Best Local Similarity 95.3%; Pred. No. 2.7e-105;
Matches 245; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

QY 5 MGELPLDINIOEPRWDSTFLGRARHFFVTYDPRNLLSGAQLASRNIVQNTAGVTP 64
DB 1 MGELPLDINIOEPRWDSTFLGRARHFFVTYDPRNLLSGAQLASRNIVQNTAGVTP 60
QY 65 GLEDDQMRKAYVYDSAFHPDGTGEKVLIGRMSAQPMMNTITGMLTFYRKTPTVVFQ 124
DB 61 GLEDDQMRKAYVYDSAFHPDGTGEKVLIGRMSAQPMMNTITGMLTFYRKTPTVVFQ 120
QY 125 WVNQSFNAIVNYSRSGDPTITVQGLTAVSATGAVATAGLKSITKGLPPLVGRFVP 184
DB 121 WVNQSFNAIVNYSRSGDPTITVQGLTAVSATGAVATAGLKSITKGLPPLVGRFVP 180
QY 185 PAAVAAANCINIPLMRQRELOVGIPTVADAGORLGISVTAAKQIGQVVISRICMAIPAM 244
DB 181 PAAVAAANCINIPLMRQRELOVGIPTVADAGORLGISVTAAKQIGQVVISRICMAIPAM 240
QY 245 AIPPLIMDTLEKKDFLK 261
DB 241 AIPPLIMDTLEKKDFLK 257

RESULT 3
SF3X_MOUSE STANDARD; PRT; 321 AA.
AC Q9IV61; O8C1Z2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Sideroflexin 3.
GN SF3X3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1). AND TISSUE SPECIFICITY.
RA MEDLINE=21172735; PubMed=11274051;
RA Fleming M.D., Campagna D.R., Haslett J.N., Trenor C.C. III,
RA Andrews N.C.;
RT "A mutation in a mitochondrial transmembrane protein is responsible
RT for the pleiotropic hematological and skeletal phenotype of
RT flexed-tail (f/t) mice".
RL Genes Dev. 15:652-657(2001).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=22354683; PubMed=12466851;
RA Okazaki Y., Futuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Balderrelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schirml L.M., Kanpin A., Matsuda H., Batalov S., Beisak K.W.,
RA Blake J.A., Bird D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalia B., Dragani T.A., Fletcher C.F., Forrest A., Fraser K.S.,
RA Gassnerland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Gilmour S., Gutlich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedziecki R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lennard B., Lyons P.A.,
RA Meglath D.R., Maltale L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.U., Pereira G., Peeble G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,

```

RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempie C.A., Seton M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita K.,
RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hironaka-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh N., Kagawa I.,
RA Miyata A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
RL Nature 420:563-573(2002).
RN (3)
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., Mcwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kerteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Sherchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield A.S., Krzywiński M.I., Skalska U., Smalls D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Potential iron transporter.
CC -1- SUBCELLULAR LOCATION: Mitochondrial (By similarity).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q91V61-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q91V61-2; Sequence=VSP_007388;
CC Note=No experimental confirmation available;
CC -1- TISSUE SPECIFICITY: Widely expressed.
CC -1- SIMILARITY: Belongs to the sideroflexin family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF325262; AAK39430.1; -;
CC EMBL: AK089885; BAC41029.1; -;
CC EMBL: BC012208; AAK12208.1; -;
CC MGI: MGI:2157679; SIKX3.
CC InterPro: IPR004686; Mtc.
CC Pfam: PF03820; Mtc; 1.
CC ProDom: PD006986; Mtc; 1.
CC TRIPDB: TRIPDB00798; mtc; 1.
CC Transport: Iron transport; Iron; Mitochondrion; Transmembrane;
CC Alternative splicing.
CC TRANSMEM 146 164
CC TRANSMEM 174 194 POTENTIAL.
CC TRANSMEM 225 245 POTENTIAL.
CC TRANSMEM 266 286 POTENTIAL.

FT VARSPLIC 112 144 Missing (in isoform 2).
FT FT 243 /FTID=VSP_007388.
SQ SEQUENCE 321 AA; 35406 MW; 9B9816A54B3F4BC CRC64;
Query Match 94.3%; Score 1269; DB 1; Length 321;
Best local similarity 95.3%; Pred. No. 5e-105; Indels 0; Gaps 0;
Matches 245; Conservative 7; Mismatches 5;
QY 5 MGELPLDINIOEPMDQSTFLGRARHFTVDPRLNLLSGAQLBASNIVQNVAGVTP 64
DB 1 MGDLPLNINIOEPMDQSTFLGRARHFTVDPRLNLLSGEQLBASNIVQNVAGVTP 60
QY 65 GITEEDLMAKXVYDSAFHPDGEKVVILGRNSAQPMMMTTTCMLTPRRTPTVFWQ 124
DB 61 GITEEDLMAKXVYDSAFHPDGEKVVILGRNSAQPMMMTTTCMLTPRRTPTVFWQ 120
QY 125 VNQNSFNALVYNSNGDPITVROGLGTAATTAATGALSKLTKHLPPLVGRFVP 184
DB 121 VNQNSFNALVYNSNGDPITVROGLGTAATTAATGALSKLTKHLPPLVGRFVP 180
QY 185 PAAVAAANCINIPLMQRELQVGPVADAGORLGYSTPAKQIFQVVISRICMAIPAM 244
DB 181 PAAVAAANCINIPLMQRELQVGPVADAGORLGYSTPAKQIFQVVISRICMAIPAM 240
QY 245 AIPPLINDLEKKDPLK 261
DB 241 AIPPLINDLEKKDPLK 257
RESULT 4
ID SFX1_HUMAN STANDARD; PRT; 322 AA.
AC Q9H9B4; Q9H9B3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Sideroflexin 1.
GN SFXN1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Li N., Chen T., Wan T., Zhang W., Cao X.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland, and Teratocarcinoma;
RA Isegai T., Ota T., Hayaishi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Wagaesuna M., Hosokiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y.,
RA Niimiya K., Iwayanagi T.;
RT "NEO human cDNA sequencing project";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 19-322 FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932;
RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., Mcwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

```
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Bickesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmitz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Might be involved in the transport of a component
CC required for iron utilization into or out of the mitochondria.
CC -1- SUBCELLULAR LOCATION: Mitochondrial (By similarity).
CC -1- SIMILARITY: Belongs to the sideroflexin family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; AF37346; AAL56007.1; -
DR EMBL; AK022287; BAB14003.1; ALT_SEQ.
DR EMBL; AK022938; BAB14318.1; ALT_INIT.
DR EMBL; BC020517; AAB20517.1; ALT_INIT.
DR Genew; HGNC:16085; SFXNL.
DR InterPro; IPR004686; Mcc.
DR Pfam; PF03820; Mcc; 1.
DR Prodom; PD006986; Mcc; 1.
DR TIGRFAMs; TIGR00798; mcc; 1.
KW Transport; Iron transport; Iron; Mitochondrion; Transmembrane.
FT TRANSMEM 103 120 POTENTIAL.
FT TRANSMEM 147 167 POTENTIAL.
FT TRANSMEM 175 195 POTENTIAL.
FT TRANSMEM 229 249 POTENTIAL.
FT TRANSMEM 267 287 POTENTIAL.
FT TRANSMEM 322 342 POTENTIAL.
SQ SEQUENCE 322 AA; 35619 MW; 47E03172P27990DB CRC64;
Query Match 81.0%; Score 1089; DB 1; Length 322;
Best Local Similarity 78.5%; Pred. No. 4,3e-89;
Matches 201; Conservative 28; Mismatches 27; Indels 0; Gaps 0;
QY 6 GELPLDINIOEPRDOSTFLGRARFFVTVDPRNLLSGAOLEASRNIVONYRAGVTPG 65
DB 3 GELPNINIKERPRDOSTFLGRARFFVTVDPRNLLSGAOLEASRNIVONYRAGVTPG 62
QY 66 ITEDOLMRARAYVDSAPHPDGEKVVLLGRMSAQVPMNMITTGCMLETFYRKTPTVVFQW 125
DB 63 LTNELMRARAYVDSAPHPDGEKVVLLGRMSAQVPMNMITTGCMLETFYRKTPTVVFQW 122
QY 126 VNOSFNALIVNSNGSDPTITVRQLGTAYVSATTGAVALGKSLTKHLPLVGRVVF 185
DB 123 INOSFNALIVNSNGSDPTITVRQLGTAYVSATTGAVALGKSLTKHLPLVGRVVF 182
QY 186 AAVAAANCINIPLMRQRELQVGIPIVADAGORLGYSVTAAKQGIQVVISIKCAIPMA 245
DB 183 AAVAAANCINIPLMRQRELQVGIPIVADAGORLGYSVTAAKQGIQVVISIKCAIPMA 242
QY 246 IPPILMDLTKKDFLK 261
DB 243 IPPILMDLTKKDFLK 258
```

```
ID SFX1_RAT STANDARD; PRT; 322 AA.
AC Q63955.
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Sideroflexin 1 (Tricarboxylate carrier protein).
GN SFXNL.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=94179133; PubMed=8132491.
RA Azzi A., Gierum M., Koller R., Mertens W., Spycher S.;
RT "The mitochondrial tricarboxylate carrier";
RL J. Bioenerg. Biomembr. 25:515-524(1993).
CC -1- FUNCTION: Might be involved in the transport of a component
CC required for iron utilization into or out of the mitochondria.
CC -1- SUBCELLULAR LOCATION: Mitochondrial.
CC -1- TISSUE SPECIFICITY: Expressed in liver and brain.
CC -1- SIMILARITY: Belongs to the sideroflexin family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; S70011; AAB30258.1; ALT_INIT.
DR InterPro; IPR004686; Mcc.
DR Pfam; PF03820; Mcc; 1.
DR Prodom; PD006986; Mcc; 1.
DR TIGRFAMs; TIGR00798; mcc; 1.
KW Transport; Iron transport; Iron; Mitochondrion; Transmembrane.
FT TRANSMEM 103 120 POTENTIAL.
FT TRANSMEM 147 167 POTENTIAL.
FT TRANSMEM 175 195 POTENTIAL.
FT TRANSMEM 229 249 POTENTIAL.
FT TRANSMEM 267 287 POTENTIAL.
SQ SEQUENCE 322 AA; 35546 MW; D28C8D89B8ABCSD CRC64;
Query Match 79.1%; Score 1064; DB 1; Length 322;
Best Local Similarity 77.3%; Pred. No. 7.1e-87;
Matches 198; Conservative 27; Mismatches 31; Indels 0; Gaps 0;
QY 6 GELPLDINIOEPRDOSTFLGRARFFVTVDPRNLLSGAOLEASRNIVONYRAGVTPG 65
DB 3 GELPNINIKERPRDOSTFLGRARFFVTVDPRNLLSGAOLEASRNIVONYRAGVTPG 62
QY 66 ITEDOLMRARAYVDSAPHPDGEKVVLLGRMSAQVPMNMITTGCMLETFYRKTPTVVFQW 125
DB 63 LTNELMRARAYVDSAPHPDGEKVVLLGRMSAQVPMNMITTGCMLETFYRKTPTVVFQW 122
QY 126 VNOSFNALIVNSNGSDPTITVRQLGTAYVSATTGAVALGKSLTKHLPLVGRVVF 185
DB 123 INOSFNALIVNSNGSDPTITVRQLGTAYVSATTGAVALGKSLTKHLPLVGRVVF 182
QY 186 AAVAAANCINIPLMRQRELQVGIPIVADAGORLGYSVTAAKQGIQVVISIKCAIPMA 245
DB 183 AAVAAANCINIPLMRQRELQVGIPIVADAGORLGYSVTAAKQGIQVVISIKCAIPMA 242
QY 246 IPPILMDLTKKDFLK 261
DB 243 IPPILMDLTKKDFLK 258
```

RESULT 5
SFX1_RATRESULT 6
SFX1_MOUSE
ID SFX1_MOUSE STANDARD; PRT; 322 AA.

AC 099RL: 09CZG4;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Sideroflexin 1.
 GN SFXN1 OR F.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND DISASE.
 RX MEDLINE=21172735; PubMed=11274051;
 RA Fleming M.D., Campagna D.R., Haslett J.N., Tremor C.C. III,
 RA Andrews N.C.;
 RT "A mutation in a mitochondrial transmembrane protein is responsible
 RT for the pleiotropic hematological and skeletal phenotype of
 RT flexed-tail (f/f) mice";
 RL Genes Dev. 15:652-657(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Embryo;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishi Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochia H.,
 RA Kuehl P., Lewis S., Maturo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baladrelli R., Barsh G.,
 RA Blake J., Boileau D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Guernicich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Membrillo P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyokawa K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohzuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection";
 RL Nature 409:685-690(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Breast;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buelow K.H., Schaefer C.F., Blat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Martusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stopleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McWeen P.J., McKernan K.J., Malek J.A., Gunatirane P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kerteman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP FUNCTION: Might be involved in the transport of a component
 RP required for iron utilization into or out of the mitochondria.
 CC -1 SUBCELLULAR LOCATION: Mitochondrial.
 CC -1 TISSUE SPECIFICITY: Widely expressed, with highest expression in
 CC kidney and liver.
 CC -1 DEVELOPMENTAL STAGE: Very high levels in the liver during the

CC period of embryonic hepatic hemopoiesis.
 CC -1 DISEASE: Defects in SFXN1 are the cause of a transitory
 CC hypochromic, microcytic anemia characterized by a large number of
 CC siderocytes containing non-heme iron granules. The anemia begins
 CC at 12 dpc, is most intense at 15 dpc and is still severe at birth,
 CC but disappears by 2 weeks of age. Mutant adults are no longer
 CC anemic, but they have an impaired response to hemopoietic stress.
 CC Most homozygotes also have flexed tails and a belly spot.
 CC -1 SIMILARITY: Belongs to the sideroflexin family.
 CC -1 CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
 CC FRAMESHIFT IN POSITION 262.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@isb-sdb.ch).
 CC -----
 CC EMBL: AF325260; AAK39428.1; -
 CC EMBL: AK012650; -; NOT ANNOTATED_CDS.
 CC EMBL: BC005743; AA05743.1; -
 CC MGD: MGI:2137677; Sfxn1.
 CC DR GO: GO:0005739; C:mitochondrion; IDA.
 CC DR GO: GO:0030218; P:erythrocyte differentiation; IMP.
 CC DR GO: GO:0006826; P:iron ion transport; IMP.
 CC DR InterPro: IPR004686; MTC.
 CC DR Pfam: PF03820; Mtc; 1.
 CC DR ProDom: PD006986; Mtc; 1.
 CC DR TrEMBL: T1G8O0798; mtc; 1.
 CC KW Transport; Iron transport; Iron; Mitochondrion; Transmembrane.
 CC FT TRANSMEM 103 120
 CC FT TRANSMEM 147 167
 CC FT TRANSMEM 175 195
 CC FT TRANSMEM 229 249
 CC FT TRANSMEM 267 287
 CC FT TRANSMEM 322 AA; 35649 MW; E3B055CB03CEDFA7 CRC64;
 CC SQ SEQUENCE
 CC Query Match 78.7%; Score 1059; DB 1; Length 322;
 CC Best Local Similarity 77.3%; Pred. No. 2e-86;
 CC Matches 198; Conservative 26; Mismatches 32; Indels 0; Gaps 0;
 CC
 CC QY 6 GELPLDINIOEPWDSTFLGRARHFFVTDPNNLLSGAQLASRNIVQYRAGVTPG 65
 CC DB 3 GEVPPNNIKKPRDQSTFFIGRAHPFTVTDPRNILLNEQLENARKVVDHYRQGIYPAG 62
 CC QY 66 ITBQQLPRAKIVYSATHTPDGKRVVLIGKMSAQVPMNMTTTCMLTFRYKTPVVFQW 125
 CC DB 63 LTNELMLRAKAYVDSAFHPDGTGEKWTILIGRMSAQVPMNMTTTCMPTFRTPPAVLFWQ 122
 CC QY 126 VNQSFNAIVVNSNGDPPIVROLGTAIVYVATGAVATLGLSKLTGHPPLVGRFVPF 185
 CC DB 123 INQSFNAIVVNTNSGDAPELVNLTGTAIVYVATGAVATLGLSKLTGHPPLVGRFVPF 182
 CC QY 186 AAVAANCINIPLRQRELQVIGIPVADAGQRLQVSYVPAKQIFQVVISICMAIPMA 245
 CC DB 183 AAVAANCINIPLRQRELKQIGIVTDENGRLQVSTPAQAQALTVQVISHILMAAPGMA 242
 CC QY 246 ITPPLIMDTLEKQDPLK 261
 CC DB 243 ITPPLIMDTLEKQDPLK 258
 CC
 CC RESULT 7
 CC SFX2_MOUSE
 CC ID SFX2_MOUSE STANDARD; PRT; 322 AA.
 CC AC Q925N2;
 CC DT 28-FEB-2003 (Rel. 41, Created)
 CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
 CC DT 15-SEP-2003 (Rel. 42, Last annotation update)
 CC DE Sideroflexin 2.
 CC GN SFXN2.

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxId=10090;
 [1]
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RX MEDLINE=21172735; PubMed=11274051;
 RA Fleming M.D., Campagna D.R., Haelet J.N., Tenor C.C. III,
 RA Andrews N.C.;
 RT "A mutation in a mitochondrial transmembrane protein is responsible
 RT for the pleiotropic hematological and skeletal phenotype of
 RT flexed-tail (f/f) mice."
 RL Genes Dev. 15:652-657(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RP TISSUE=Kidney;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulhaly S.J.,
 RA Bosak S.A., Mcswan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W.,
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalek U., Smalins D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences".
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC - FUNCTION: Potential Iron transporter.
 CC - SUBCELLULAR LOCATION: Mitochondrial (By similarity).
 CC - TISSUE SPECIFICITY: Expressed in brain, heart, kidney, spleen,
 CC thymus, liver, stomach and skin.
 CC - SIMILARITY: Belongs to the sideroflexin family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.ebi.ac.uk/announcements/](http://www.ebi.ac.uk/announcements)
 CC or send an email to license@ebi.ac.uk).
 CC -----
 CC DR EMBL; AF325261; AA039429.1; -;
 CC DR EMBL; BC019808; AA019808.1; -;
 CC DR MGI; MGI:2137678; Sfxn2.
 CC InterPro: IPR004686; MTC.
 CC Pfam: PF03820; MTC; 1.
 CC DR Pfam; PF03820; MTC; 1.
 CC DR PRODOM; PD006986; MTC; 1.
 CC DR TIGRFAMs; TIGR00798; MTC; 1.
 CC KM Transprot; Iron transporter; Iron; Mitochondrion; Transmembrane.
 FT TRANSMEM 99 119 POTENTIAL.
 FT TRANSMEM 147 167 POTENTIAL.
 FT TRANSMEM 174 194 POTENTIAL.
 FT TRANSMEM 223 243 POTENTIAL.
 FT TRANSMEM 266 286 POTENTIAL.
 FT CONFLICT 3 3 G -> A (IN REF. 2).
 SO SEQUENCE 322 AA; 36141 MW; 00D6236898B983C CRC64;

DB 9 NIDAPRWDCTFLGRVYHFNITDPTVPASFOELDMAVAIVKSMGLVPTQMEQL 68
 QY 73 RAKYVYDASAFHDPDGEKVLIGRMSAQVPMNNTTTCMLTPRKTPTVVPWQVNSFNA 132
 DB 69 YAKKIDYSAFHPDGEKKNVIGRMSFQVPGCMILITFPMLOFYRTMPAVLFWQVNSFNA 128
 QY 133 IYVNSRSDPTLTROLOTAIVASATGAVATPLGLKSLTKPLPLVGRFVPPAAVAAAN 192
 DB 129 LVNYTRNASPSVQAQMLSYFATTTTAVATVAGNMMWKAPPLVGRWVPPAAVAAAN 188
 QY 193 CINIPLMROREOVGIPVADGQRGLGYSVTAAGKQIPQVVISRICMAIPMAIPPLMD 252
 DB 189 CNIIPMRQDELQICVDRNQLGHSQRAAAVGIAGVVISRIMAPGMLLLEVINE 248
 QY 253 TLEKDFLK 261
 DB 249 RLRLHLMK 257
 RESULT 8
 ID SFX2_HUMAN STANDARD; PRT; 322 AA.
 AC 096NB2;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Sideroflexin 2.
 GN SFXN2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RP TISSUE=Muscle;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulhaly S.J.,
 RA Bosak S.A., Mcswan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W.,
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalek U., Smalins D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences".
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE OF 1-204 FROM N.A.
 RP TISSUE=Neuroblastoma;
 RC Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
 RA Fukumori T., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
 RA Otsuki T., Sato H., Wakamatsu A., Ichii S., Yamamoto Y., Isono Y.,
 RA Kawai-Hiro Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
 RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda A., Nagatsuma M.,
 RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
 RA Kawakami B., Suzuki Y., Sugano S., Nagahart K., Masuno Y., Nagai K.,
 RA Isogai T.;
 RT "NEDO human cDNA sequencing project".
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 CC - FUNCTION: Potential Iron transporter.
 CC - SUBCELLULAR LOCATION: Mitochondrial (By similarity).
 CC - SIMILARITY: Belongs to the sideroflexin family.


```

CC      or send an email to license@lsb-sib.ch).
CC
CC      EMBL, AY044437, XAN95826.1; -.
CC      Genew; HGNC:16073; SFXNS.
CC      InterPro; IPR004686; Mtc.
CC      Pfam; PF03820; Mtc; 1.
CC      ProDom; PD009686; Mtc; 1.
CC      Trifam; TRF00798; mtc; 1.
CC      Transports; Iron transport; Iron; Mitochondrion; Transmembrane.
CC      Transmem 103 123 POTENTIAL.
CC      Transmem 163 183 POTENTIAL.
CC      Transmem 254 274 POTENTIAL.
CC      Transmem 287 307 POTENTIAL.
CC      SEQUENCE 340 AA; 37124 MW; AF4103C7DB3F2DAC CEC64;

```

Query March	36.0%	Score 484.5	DB 1	Length 340
Best Local Similarity	37.9%	Pred. No. 1.5e-35		
Matches 96	Conservative 51	Mismatches 99	Indels 7	Gaps 2

[illegible]

RESULT 11	SEFX_MOUSE	STANDARD	PRT	342 AA.
ID	SEFX_MOUSE			
AC	Q925N0; QBEGG3;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	15-SEP-2003 (Rel. 42, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Sideroflexin 5.			
GN	SEFXN5.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
NCBI	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N. A. (ISOFORM 2).			
RX	MEDLINE=22354683; PubMed=12466851;			
RA	Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,			
RA	Nikaido I., Oeato N., Saito R., Suzuki H., Yamanae I., Kiyosawa H.,			
RA	Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gotohori T.,			
RA	Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,			
RA	Schirral L.M., Knapin A., Matveeva H., Batalov S., Beisel K.W.,			
RA	Blake J.A., Bradt D., Bruscia V., Chochoia C., Corbali L.E., Cousins S.,			
RA	Dalla E., Driagani T.A., Fletcher C.F., Forrest A., Frazier K.S.,			
RA	Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,			
RA	Grimmond S., Guslichich S., Hitokawa N., Jackson I.J., Jarvis E.D.,			
RA	Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,			
RA	Konagaya A., Kurochkin I.V., Lee Y., Lennard B., Lyons P.A.,			
RA	Majloft D.R., Matsuda K., Marchionni L., McKenzie L., Miki H.,			
RA	Nagasshima T., Numa K., Okido T., Pavan W.J., Petrea G., Pesole G.,			

RA Petrovsky N., Pillai R., Pontius J.V., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sendelin A., Schneider C., Sempke C.A., Setou M., Shimada K.,
RA Sultana R., Takenawa Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verrardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wymshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Cerniack P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kigawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Wacrezon R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.",
RL Nature 420:563-573(2002).

RP SHOUNCNC F 1-181 FROM N.A., AND TISSUE SPECIFICITY.
RX MEDLINE=21172735; PubMed=11274051;
RA Fleming M.D., Campagna D.R., Haelet J.N., Trenor C.C. III,
RA Andrews N.C.;
RT "A mutation in a mitochondrial transmembrane protein is responsible
RT for the pleiotropic hematological and skeletal phenotype of
RT flexed-tail (E/E) mice."
RL Genes Dev. 15:652-657(2001).
CC -I- FUNCTION: Potential iron transporter.
CC -I- SUBCELLULAR LOCATION: Mitochondrial (By similarity).
CC -I- TISSUE SPECIFICITY: Expressed in liver and brain.
CC -I- SIMILARITY: Belongs to the sideroflexin family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation --
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and that statement is not removed. Usage by and for commercial
CC entities requires a licenseesib-sib.ch).
CC or send an email to license@sib-sib.ch.
CC

DR	EMBL; AK036285;	BAC29372.1;	-;
DR	EMBL; AK082073;	BAC38403.1;	-;
DR	EMBL; AF325264;	AAC39432.1;	-;
DR	MGD; MGI:2137681;	Sixns.	
DR	InterPro; IPR004686;	Mtc.	
	Pfam; PF03820;	Mtc.	1.
KM	Transport; Iron transport; Iron; Mitochondrion; Transmembrane.		
FT	TRANSMEM	105	125
FT	TRANSMEM	165	185
FT	TRANSMEM	256	276
FT	TRANSMEM	289	309
QO	SEQUENCE	342 AA;	37328 MW; 9BDB8261D5EF79D9 CRC64;

```

Query Match Score 484.5; DB 1; Length 342;
Best Local Similarity 38.3%; Pred. No. 1.5e-35;
Matches 97; Conservative 50; Mismatches 99; Indels 7; Gaps 2;

```

QY 16 EPPNDOSTEJAKAHKEFTVIDPKLLLSGAOLBSKAI VONIKKGVVIFGLIBOLMKA 75
 Db 31 KPRFOQTSFYGRFHEFLDIDPRTLFTXEKRLREAVOLLBRYKKGITLRPGVTNQLMSAQ 90
 QY 76 YVVDASHPPTGEXVVLIGRMSAQVPMNMITGCMLEFRKTPVFPVQMNOSFNAIVN 135
 Db 91 KIKOAILHPNTNEKIFMPFMSGIIPEPGTPIVVELLPNQTLASTVFMQIMOSHMCYN 150
 QY 136 YSNRSGDPTTVQQLGTAIVYSAITGAVALGLKSL---TKHLRP---LVGRFVPPAAV 188
 Db 151 YARNNAKPEBPASKEFIGYGAIVSAVSIVGLNVLVQKANKFEPATRLVQRFVPPAV 210
 QY 189 AANNCITIPMBROKELQVGI PVADBAQORGYSTAKOGIFQVIVSRICAI PAMAIP 248
 Db 211 ASANICNVMLRGELEEGIDVDADQNLVGYSSKIARHALLERATLRVVLPMPIVLP 270
 QY 249 LIMDTIEKKDFLK 261

Db 271 IYMSMLEKTAALQ 283

RESULT 12

ID	SRE2 CAEEL	STANDARD;	PRT;	600 AA.
AC	009273;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	15-SEP-2003 (Rel. 42, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DN	Serpentine receptor class epsilon 2 (Sre-2 protein).			
GN	SRE-2 OR C41C.2.			
OS	Caenorhabditis elegans.			
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;			
CC	Rhabditidae; Peloderinae; Caenorhabditis.			
OK	NCBI_Taxid=6239;			
RN	SEQUENCE FROM N.A.			
RP	STRAIN=Briscot N2;			
RA	Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.			
RN	Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.			
RP	REVISIONS.			
RA	Dustin R.;			
RL	Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein (potential).			
CC	-1- SIMILARITY: BELONGS TO THE C.ELEGANS RECEPTOR-LIKE PROTEIN SRE FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; Z48045; CAAB8099.2; -			
DR	PIR; T19873; T19873.			
DR	WormPep; C41C.2; CE32323.			
DR	InterPro; IPR004686; Mtc.			
DR	InterPro; IPR004151; Sre.			
DR	Pfam; PF03820; Mtc; 1.			
DR	Pfam; PF03125; Sre; 1.			
DR	Prodom; PD006986; Mtc; 1.			
DR	TIGRFAMs; TIGR00798; mtc; 1.			
KW	Transmembrane; Multigene family.			
FT	TRANSMEM 129 149 POTENTIAL.			
FT	TRANSMEM 178 198 POTENTIAL.			
FT	TRANSMEM 269 289 POTENTIAL.			
FT	TRANSMEM 303 323 POTENTIAL.			
FT	TRANSMEM 380 400 POTENTIAL.			
FT	TRANSMEM 410 430 POTENTIAL.			
FT	TRANSMEM 466 486 POTENTIAL.			
FT	TRANSMEM 506 526 POTENTIAL.			
SO	SEQUENCE 600 AA; 68517 MW; 418AC8FB34C44CDS CRC64;			

Query Match 32.2%; Score 432.5; DB 1; Length 600;
 Best Local Similarity 40.0%; Pred. No. 1.2e-30;
 Matches 100; Conservative 43; Mismatches 96; Indels 11; Gaps 6;

Qy 188 VAAANCINIPIMROELGVIPVAD-EAGORUGSVTAAKOGIPQVVISRICMAIPAMAI 246

Db 222 TSLASLVVICRKNNELESTGIVYEKDTGVGVSKVAKAQVDTTWRAPVPLILM 281

Qy 247 PPLIMDTLEK 256

Db 282 PRCIMPYLER 291

RESULT 13

ID	SFX4 MOUSE	STANDARD;	PRT;	313 AA.
AC	0925N1;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DN	Sideroflexin 4.			
GN	SFXN4.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
CC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OK	NCBI_Taxid=10090;			
RN	SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.			
RP	Medline=21172735; PubMed=11274051;			
RA	Fleming M.D., Campagna D.R., Haslett J.N., Trenor C.C. III, Andrews N.C.;			
RT	"A mutation in a mitochondrial transmembrane protein is responsible for the pleiotropic hematological and skeletal phenotype of flexed-tail (f/f) mice."			
RL	Genes Dev. 15:652-657(2001).			
CC	-1- FUNCTION: Potential iron transporter.			
CC	-1- SUBCELLULAR LOCATION: Mitochondrial (By similarity).			
CC	-1- TISSUE SPECIFICITY: Largely restricted to kidney, brain and heart.			
CC	-1- SIMILARITY: Belongs to the sideroflexin family.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; AF325263; AAK39431.1; -			
DR	MGI; MGI:2137680; Sfxn4.			
DR	InterPro; IPR004686; Mtc.			
DR	Pfam; PF03820; Mtc; 1.			
DR	Prodom; PD006986; Mtc; 1.			
KW	Transport; Iron transport; Iron; Mitochondrion; Transmembrane.			
FT	TRANSMEM 87 107 POTENTIAL.			
FT	TRANSMEM 141 161 POTENTIAL.			
FT	TRANSMEM 175 191 POTENTIAL.			
FT	TRANSMEM 230 247 POTENTIAL.			
FT	TRANSMEM 269 289 POTENTIAL.			
SO	SEQUENCE 313 AA; 35685 MW; ACP650B711A7552P CRC64;			

Query Match 8.4%; Score 113; DB 1; Length 313;
 Best Local Similarity 20.7%; Pred. No. 0.011;
 Matches 56; Conservative 50; Mismatches 125; Indels 40; Gaps 8;

QY 173 KHLPL- ---VGFVPPAAVAANINIPMRQRELQVGVADDEAGQRLGYSVTAAKQ 228
 DB 166 ----PLNNFWLKRRLPLVFLAQVSGMNVFASRSEFNHRIEIVMDKGVGHGRKGRKA 221
 QY 229 IFQVVISRICMAIPMAIPLIMDTLEKDF 259
 DB 222 IKDTAKSRAVLFGTSLAPLFLFIHFRTRF 252

RESULT 14
 GCSR_HUMAN STANDARD; PRT; 836 AA.
 AC 099062;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Granulocyte colony stimulating factor receptor precursor (G-CSF-R)
 DE (CD114 antigen).
 GN CSF3R OR GCSFR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=91079757; PubMed=2147944;
 RA Larsen A., Davis T., Curtis B.M., Gimpel S., Sims J.E., Cosman D.,
 RT "Expression cloning of a human granulocyte colony-stimulating factor
 RT receptor: a structural mosaic of hematopoietin receptor,
 RT immunoglobulin, and fibronectin domains.";
 RL J. Exp. Med. 172:1559-1570 (1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=91062348; PubMed=1701053;
 RA Fukunaga R., Seto Y., Mizushima S., Nagata S.;
 RT "Three different mRNAs encoding human granulocyte colony-stimulating
 RT factor receptor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:8702-8706 (1990).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92091782; PubMed=1530796;
 RA Seto Y., Fukunaga R., Nagata S.;
 RT "Chromosomal gene organization of the human granulocyte colony-
 RT stimulating factor receptor.";
 RL J. Immunol. 148:259-266 (1992).
 RN [4]
 RP SEQUENCE FROM N.A. AND VARIANTS THR-231; ASN-320; ARG-346; LYS-405;
 RP GIN-440; HIS-510; HIS-562 AND CYS-583.
 RA Rieder M.J., Armel T.Z., Carrington D.P., Ozuna M., Kuldanek S.A.,
 RA Rajkumar N., Toch E.J., Yi O., Nickerson D.A.,
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP DOMAINS STRUCTURE.
 RX MEDLINE=92007729; PubMed=1717255;
 RA Fukunaga R., Ishizaka-Ikeda E., Pan C.-X., Seto Y., Nagata S.;
 RT "Functional domains of the granulocyte colony-stimulating factor
 RT receptor.";
 RL EMBO J. 10:2855-2865 (1991).
 RN [6]
 RP DISEASE.
 RX MEDLINE=94240159; PubMed=7514305;
 RA Dong F., Hoefleot L.H., Schelen A.M., Broeders C.A., Melijer Y.,
 RA Verma A.J., Tow I.P., Lowenberg B.,
 RT "Identification of a nonsense mutation in the granulocyte-colony-
 RT stimulating factor receptor in severe congenital neutropenia.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:4480-4484 (1994).
 RN [7]
 RP STRUCTURE BY NMR OF 227-334.
 RX MEDLINE=97331327; PubMed=187659;
 RA Yamaeaki K., Naito S., Anaguchi H., Ohkubo T., Ota Y.;

RT "Solution structure of an extracellular domain containing the WSxWS
 RT motif of the granulocyte colony-stimulating factor receptor and its
 RT interaction with ligand.";
 RL Nat. Struct. Biol. 4:498-503 (1997).
 RN [8]
 RP 3D-STRUCTURE MODELING OF 125-331.
 RX MEDLINE=98037802; PubMed=9368043;
 RA Layton J.E., Iaria J., Smith D.K., Treutlein H.R.;
 RT "Identification of a ligand-binding site on the granulocyte colony-
 RT stimulating factor receptor by molecular modeling and mutagenesis.";
 RL J. Biol. Chem. 272:29735-29741 (1997).
 CC -1- FUNCTION: RECEPTOR FOR GRANULOCYTE COLONY-STIMULATING FACTOR (G-
 CC CSF). IN ADDITION IT MAY FUNCTION IN SOME ADHESION OR RECOGNITION
 CC EVENTS AT THE CELL SURFACE.
 CC -1- SUBUNIT: DIMER (PROBABLE).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. THE GCSFR-2 FORM,
 CC WHICH LACKS THE TRANSMEMBRANE DOMAIN, MAY REPRESENT A SOLUBLE FORM
 CC OF THE RECEPTOR.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=4;
 CC Comment=Additional isoforms seem to exist. Experimental
 CC confirmation may be lacking for some isoforms;
 CC Name=1; Synonyms=GCSFR-1; Sequence=Displayed;
 CC IsoId=Q99062-1; Sequence=VSP_001674;
 CC Name=2; Synonyms=GCSFR-2;
 CC IsoId=Q99062-2; Sequence=VSP_001674;
 CC Name=3; Synonyms=GCSFR-3;
 CC IsoId=Q99062-3; Sequence=VSP_001673;
 CC Name=4; Synonyms=GCSFR-4, D7;
 CC IsoId=Q99062-4; Sequence=VSP_001671, VSP_001672;
 CC -1- TISSUE SPECIFICITY: ONE OR SEVERAL FORMS HAVE BEEN FOUND IN
 CC MYELOGENOUS LEUKEMIA CELL LINE KG-1, LEUKEMIA U937 CELL LINE, IN
 CC BONE MARROW CELLS, PLACENTA, AND PERIPHERAL BLOOD GRANULOCYTES. THE
 CC GCSFR-2 FORM HAS BEEN FOUND ONLY IN LEUKEMIA U937 CELLS. THE
 CC GCSFR-3 FORM IS HIGHLY EXPRESSED IN PLACENTA.
 CC -1- DISEASE: Defects in CSF3R might be a cause of severe congenital
 CC neutropenia (SCN) in some patients.
 CC -1- SIMILARITY: Contains 5 fibronectin type III domains.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC -1- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.
 CC -1- DATABASE: NAME=PROM; NOTE=CD guide CD14 entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd114.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X55721; CAA39253.1; -
 DR EMBL: X55720; CAA39252.1; -
 DR EMBL: S71484; AAB20660.1; -
 DR EMBL: M59818; AAA63176.1; -
 DR EMBL: M59819; AAA63177.1; -
 DR EMBL: M59820; AAA63178.1; -
 DR EMBL: AY148100; AAN05790.1; -
 DR PIR: B38252; B38252.
 DR PIR: C38252; C38252.
 DR PIR: C38253; C38253.
 DR PIR: C38254; C38254.
 DR PIR: C38255; C38255.
 DR PIR: C38256; C38256.
 DR PIR: C38257; C38257.
 DR PIR: C38258; C38258.
 DR PIR: C38259; C38259.
 DR PIR: C38260; C38260.
 DR PIR: C38261; C38261.
 DR PIR: C38262; C38262.
 DR PIR: C38263; C38263.
 DR PIR: C38264; C38264.
 DR PIR: C38265; C38265.
 DR PIR: C38266; C38266.
 DR PIR: C38267; C38267.
 DR PIR: C38268; C38268.
 DR PIR: C38269; C38269.
 DR PIR: C38270; C38270.
 DR PIR: C38271; C38271.
 DR PIR: C38272; C38272.
 DR PIR: C38273; C38273.
 DR PIR: C38274; C38274.
 DR PIR: C38275; C38275.
 DR PIR: C38276; C38276.
 DR PIR: C38277; C38277.
 DR PIR: C38278; C38278.
 DR PIR: C38279; C38279.
 DR PIR: C38280; C38280.
 DR PIR: C38281; C38281.
 DR PIR: C38282; C38282.
 DR PIR: C38283; C38283.
 DR PIR: C38284; C38284.
 DR PIR: C38285; C38285.
 DR PIR: C38286; C38286.
 DR PIR: C38287; C38287.
 DR PIR: C38288; C38288.
 DR PIR: C38289; C38289.
 DR PIR: C38290; C38290.
 DR PIR: C38291; C38291.
 DR PIR: C38292; C38292.
 DR PIR: C38293; C38293.
 DR PIR: C38294; C38294.
 DR PIR: C38295; C38295.
 DR PIR: C38296; C38296.
 DR PIR: C38297; C38297.
 DR PIR: C38298; C38298.
 DR PIR: C38299; C38299.
 DR PIR: C38300; C38300.
 DR PIR: C38301; C38301.
 DR PIR: C38302; C38302.
 DR PIR: C38303; C38303.
 DR PIR: C38304; C38304.
 DR PIR: C38305; C38305.
 DR PIR: C38306; C38306.
 DR PIR: C38307; C38307.
 DR PIR: C38308; C38308.
 DR PIR: C38309; C38309.
 DR PIR: C38310; C38310.
 DR PIR: C38311; C38311.
 DR PIR: C38312; C38312.
 DR PIR: C38313; C38313.
 DR PIR: C38314; C38314.
 DR PIR: C38315; C38315.
 DR PIR: C38316; C38316.
 DR PIR: C38317; C38317.
 DR PIR: C38318; C38318.
 DR PIR: C38319; C38319.
 DR PIR: C38320; C38320.
 DR PIR: C38321; C38321.
 DR PIR: C38322; C38322.
 DR PIR: C38323; C38323.
 DR PIR: C38324; C38324.
 DR PIR: C38325; C38325.
 DR PIR: C38326; C38326.
 DR PIR: C38327; C38327.
 DR PIR: C38328; C38328.
 DR PIR: C38329; C38329.
 DR PIR: C38330; C38330.
 DR PIR: C38331; C38331.
 DR PIR: C38332; C38332.
 DR PIR: C38333; C38333.
 DR PIR: C38334; C38334.
 DR PIR: C38335; C38335.
 DR PIR: C38336; C38336.
 DR PIR: C38337; C38337.
 DR PIR: C38338; C38338.
 DR PIR: C38339; C38339.
 DR PIR: C38340; C38340.
 DR PIR: C38341; C38341.
 DR PIR: C38342; C38342.
 DR PIR: C38343; C38343.
 DR PIR: C38344; C38344.
 DR PIR: C38345; C38345.
 DR PIR: C38346; C38346.
 DR PIR: C38347; C38347.
 DR PIR: C38348; C38348.
 DR PIR: C38349; C38349.
 DR PIR: C38350; C38350.
 DR PIR: C38351; C38351.
 DR PIR: C38352; C38352.
 DR PIR: C38353; C38353.
 DR PIR: C38354; C38354.
 DR PIR: C38355; C38355.
 DR PIR: C38356; C38356.
 DR PIR: C38357; C38357.
 DR PIR: C38358; C38358.
 DR PIR: C38359; C38359.
 DR PIR: C38360; C38360.
 DR PIR: C38361; C38361.
 DR PIR: C38362; C38362.
 DR PIR: C38363; C38363.
 DR PIR: C38364; C38364.
 DR PIR: C38365; C38365.
 DR PIR: C38366; C38366.
 DR PIR: C38367; C38367.
 DR PIR: C38368; C38368.
 DR PIR: C38369; C38369.
 DR PIR: C38370; C38370.
 DR PIR: C38371; C38371.
 DR PIR: C38372; C38372.
 DR PIR: C38373; C38373.
 DR PIR: C38374; C38374.
 DR PIR: C38375; C38375.
 DR PIR: C38376; C38376.
 DR PIR: C38377; C38377.
 DR PIR: C38378; C38378.
 DR PIR: C38379; C38379.
 DR PIR: C38380; C38380.
 DR PIR: C38381; C38381.
 DR PIR: C38382; C38382.
 DR PIR: C38383; C38383.
 DR PIR: C38384; C38384.
 DR PIR: C38385; C38385.
 DR PIR: C38386; C38386.
 DR PIR: C38387; C38387.
 DR PIR: C38388; C38388.
 DR PIR: C38389; C38389.
 DR PIR: C38390; C38390.
 DR PIR: C38391; C38391.
 DR PIR: C38392; C38392.
 DR PIR: C38393; C38393.
 DR PIR: C38394; C38394.
 DR PIR: C38395; C38395.
 DR PIR: C38396; C38396.
 DR PIR: C38397; C38397.
 DR PIR: C38398; C38398.
 DR PIR: C38399; C38399.
 DR PIR: C38400; C38400.
 DR PIR: C38401; C38401.
 DR PIR: C38402; C38402.
 DR PIR: C38403; C38403.
 DR PIR: C38404; C38404.
 DR PIR: C38405; C38405.
 DR PIR: C38406; C38406.
 DR PIR: C38407; C38407.
 DR PIR: C38408; C38408.
 DR PIR: C38409; C38409.
 DR PIR: C38410; C38410.
 DR PIR: C38411; C38411.
 DR PIR: C38412; C38412.
 DR PIR: C38413; C38413.
 DR PIR: C38414; C38414.
 DR PIR: C38415; C38415.
 DR PIR: C38416; C38416.
 DR PIR: C38417; C38417.
 DR PIR: C38418; C38418.
 DR PIR: C38419; C38419.
 DR PIR: C38420; C38420.
 DR PIR: C38421; C38421.
 DR PIR: C38422; C38422.
 DR PIR: C38423; C38423.
 DR PIR: C38424; C38424.
 DR PIR: C38425; C38425.
 DR PIR: C38426; C38426.
 DR PIR: C38427; C38427.
 DR PIR: C38428; C38428.
 DR PIR: C38429; C38429.
 DR PIR: C38430; C38430.
 DR PIR: C38431; C38431.
 DR PIR: C38432; C38432.
 DR PIR: C38433; C38433.
 DR PIR: C38434; C38434.
 DR PIR: C38435; C38435.
 DR PIR: C38436; C38436.
 DR PIR: C38437; C38437.
 DR PIR: C38438; C38438.
 DR PIR: C38439; C38439.
 DR PIR: C38440; C38440.
 DR PIR: C38441; C38441.
 DR PIR: C38442; C38442.
 DR PIR: C38443; C38443.
 DR PIR: C38444; C38444.
 DR PIR: C38445; C38445.
 DR PIR: C38446; C38446.
 DR PIR: C38447; C38447.
 DR PIR: C38448; C38448.
 DR PIR: C38449; C38449.
 DR PIR: C38450; C38450.
 DR PIR: C38451; C38451.
 DR PIR: C38452; C38452.
 DR PIR: C38453; C38453.
 DR PIR: C38454; C38454.
 DR PIR: C38455; C38455.
 DR PIR: C38456; C38456.
 DR PIR: C38457; C38457.
 DR PIR: C38458; C38458.
 DR PIR: C38459; C38459.
 DR PIR: C38460; C38460.
 DR PIR: C38461; C38461.
 DR PIR: C38462; C38462.
 DR PIR: C38463; C38463.
 DR PIR: C38464; C38464.
 DR PIR: C38465; C38465.
 DR PIR: C38466; C38466.
 DR PIR: C38467; C38467.
 DR PIR: C38468; C38468.
 DR PIR: C38469; C38469.
 DR PIR: C38470; C38470.
 DR PIR: C38471; C38471.
 DR PIR: C38472; C38472.
 DR PIR: C38473; C38473.
 DR PIR: C38474; C38474.
 DR PIR: C38475; C38475.
 DR PIR: C38476; C38476.
 DR PIR: C38477; C38477.
 DR PIR: C38478; C38478.
 DR PIR: C38479; C38479.
 DR PIR: C38480; C38480.
 DR PIR: C38481; C38481.
 DR PIR: C38482; C38482.
 DR PIR: C38483; C38483.
 DR PIR: C38484; C38484.
 DR PIR: C38485; C38485.
 DR PIR: C38486; C38486.
 DR PIR: C38487; C38487.
 DR PIR: C38488; C38488.
 DR PIR: C38489; C38489.
 DR PIR: C38490; C38490.
 DR PIR: C38491; C38491.
 DR PIR: C38492; C38492.
 DR PIR: C38493; C38493.
 DR PIR: C38494; C38494.
 DR PIR: C38495; C38495.
 DR PIR: C38496; C38496.
 DR PIR: C38497; C38497.
 DR PIR: C38498; C38498.
 DR PIR: C38499; C38499.
 DR PIR: C38500; C38500.
 DR PIR: C38501; C38501.
 DR PIR: C38502; C38502.
 DR PIR: C38503; C38503.
 DR PIR: C38504; C38504.
 DR PIR: C38505; C38505.
 DR PIR: C38506; C38506.
 DR PIR: C38507; C38507.
 DR PIR: C38508; C38508.
 DR PIR: C38509; C38509.
 DR PIR: C38510; C38510.
 DR PIR: C38511; C38511.
 DR PIR: C38512; C38512.
 DR PIR: C38513; C38513.
 DR PIR: C38514; C38514.
 DR PIR: C38515; C38515.
 DR PIR: C38516; C38516.
 DR PIR: C38517; C38517.
 DR PIR: C38518; C38518.
 DR PIR: C38519; C38519.
 DR PIR: C38520; C38520.
 DR PIR: C38521; C38521.
 DR PIR: C38522; C38522.
 DR PIR: C38523; C38523.
 DR PIR: C38524; C38524.
 DR PIR: C38525; C38525.
 DR PIR: C38526; C38526.
 DR PIR: C38527; C38527.
 DR PIR: C38528; C38528.
 DR PIR: C38529; C38529.
 DR PIR: C38530; C38530.
 DR PIR: C38531; C38531.
 DR PIR: C38532; C38532.
 DR PIR: C38533; C38533.
 DR PIR: C38534; C38534.
 DR PIR: C38535; C38535.
 DR PIR: C38536; C38536.
 DR PIR: C38537; C38537.
 DR PIR: C38538; C38538.
 DR PIR: C38539; C38539.
 DR PIR: C38540; C38540.
 DR PIR: C38541; C38541.
 DR PIR: C38542; C38542.
 DR PIR: C38543; C38543.
 DR PIR: C38544; C38544.
 DR PIR: C38545; C38545.
 DR PIR: C38546; C38546.
 DR PIR: C38547; C38547.
 DR PIR: C38548; C38548.
 DR PIR: C38549; C38549.
 DR PIR: C38550; C38550.
 DR PIR: C38551; C38551.
 DR PIR: C38552; C38552.
 DR PIR: C38553; C38553.
 DR PIR: C38554; C38554.
 DR PIR: C38555; C38555.
 DR PIR: C38556; C38556.
 DR PIR: C38557; C38557.
 DR PIR: C38558; C38558.
 DR PIR: C38559; C38559.
 DR PIR: C38560; C38560.
 DR PIR: C38561; C38561.
 DR PIR: C38562; C38562.
 DR PIR: C38563; C38563.
 DR PIR: C38564; C38564.
 DR PIR: C38565; C38565.
 DR PIR: C38566; C38566.
 DR PIR: C38567; C38567.
 DR PIR: C38568; C38568.
 DR PIR: C38569; C38569.
 DR PIR: C38570; C38570.
 DR PIR: C38571; C38571.
 DR PIR: C38572; C38572.
 DR PIR: C38573; C38573.
 DR PIR: C38574; C38574.
 DR PIR: C38575; C38575.
 DR PIR: C38576; C38576.
 DR PIR: C38577; C38577.
 DR PIR: C38578; C38578.
 DR PIR: C38579; C38579.
 DR PIR: C38580; C38580.
 DR PIR: C38581; C38581.
 DR PIR: C38582; C38582.
 DR PIR: C38583; C38583.
 DR PIR: C38584; C38584.
 DR PIR: C38585; C38585.
 DR PIR: C38586; C38586.
 DR PIR: C38587; C38587.
 DR PIR: C38588; C38588.
 DR PIR: C38589; C38589.
 DR PIR: C38590; C38590.
 DR PIR: C38591; C38591.
 DR PIR: C38592; C38592.
 DR PIR: C38593; C38593.
 DR PIR: C38594; C38594.
 DR PIR: C38595; C38595.
 DR PIR: C38596; C38596.
 DR PIR: C38597; C38597.
 DR PIR: C38598; C38598.
 DR PIR: C38599; C38599.
 DR PIR: C38600; C38600.
 DR PIR: C38601; C38601.
 DR PIR: C38602; C38602.
 DR PIR: C38603; C38603.
 DR PIR: C38604; C38604.
 DR PIR: C38605; C38605.
 DR PIR: C38606; C38606.
 DR PIR: C38607; C38607.
 DR PIR: C38608; C38608.
 DR PIR: C38609; C38609.
 DR PIR: C38610; C38610.
 DR PIR: C38611; C38611.
 DR PIR: C38612; C38612.
 DR PIR: C38613; C38613.
 DR PIR: C38614; C38614.
 DR PIR: C38615; C38615.
 DR PIR: C38616; C38616.
 DR PIR: C38617; C38617.
 DR PIR: C38618; C38618.
 DR PIR: C38619; C38619.
 DR PIR: C38620; C38620.
 DR PIR: C38621; C38621.
 DR PIR: C38622; C38622.
 DR PIR: C38623; C38623.
 DR PIR: C38624; C38624.
 DR PIR: C38625; C38625.
 DR PIR: C38626; C38626.
 DR PIR: C38627; C38627.
 DR PIR: C38628; C38628.
 DR PIR: C38629; C38629.
 DR PIR: C38630; C38630.
 DR PIR: C38631; C38631.
 DR PIR: C38632; C38632.
 DR PIR: C38633; C38633.
 DR PIR: C38634; C38634.
 DR PIR: C38635; C38635.
 DR PIR: C38636; C38636.
 DR PIR: C38637; C38637.
 DR PIR: C38638; C38638.
 DR PIR: C38639; C38639.
 DR PIR: C38640; C38640.
 DR PIR: C38641; C38641.
 DR PIR: C38642; C38642.
 DR PIR: C38643; C38643.
 DR PIR: C38644; C38644.
 DR PIR: C38645; C38645.
 DR PIR: C38646; C38646.
 DR PIR: C38647; C38647.
 DR PIR: C38648; C38648.
 DR PIR: C38649; C38649.
 DR PIR: C38650; C38650.
 DR PIR: C38651; C38651.
 DR PIR: C38652; C38652.
 DR PIR: C38653; C38653.
 DR PIR: C38654; C38654.
 DR PIR: C38655; C38655.
 DR PIR: C38656; C38656.
 DR PIR: C38657; C38657.
 DR PIR: C38658; C38658.
 DR PIR: C38659; C38659.
 DR PIR: C38660; C38660.
 DR PIR: C38661; C38661.
 DR PIR: C38662; C38662.
 DR PIR: C38663; C38663.
 DR PIR: C38664; C38664.
 DR PIR: C38665; C38665.
 DR PIR: C38666; C38666.
 DR PIR: C38667; C38667.
 DR PIR: C38668; C38668.
 DR PIR: C38669; C38669.
 DR PIR: C38670; C38670.
 DR PIR: C38671; C38671.
 DR PIR: C38672; C38672.
 DR PIR: C38673; C38673.
 DR PIR: C38674; C38674.
 DR PIR: C38675; C38675.
 DR PIR: C38676; C38676.
 DR PIR: C38677; C38677.
 DR PIR: C38678; C38678.
 DR PIR: C38679; C38679.
 DR PIR: C38680; C38680.
 DR PIR: C38681; C38681.<


```
FT DISULFID 243 295 POTENTIAL.
FT CARBOHYD 44 44 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 70 70 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 216 216 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 284 284 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 305 305 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 321 321 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT LIPID 321 321 GPI-ANCHOR (POTENTIAL).
SQ SEQUENCE 344 AA, 37998 MW, CBB39B853B3B224 CRC64,
```

```
Query Match. 6.7%; Score 89.5; DB 1; Length 344;
Best Local Similarity 24.8%; Pred. No. 1.4;
Matches 52; Conservative 26; Mismatches 95; Indels 35; Gaps 10;
```

```
QY 2 ESKMGELPDIDINIQEPRW-DQSTFLGRARHFTVTDPRNLLSGAQLASRNI--VQNT 58
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 51 ESATIRCTIDNRVTRVAMLRSTILYAGNDKMC-LDPRVLLSMTQTQYSIEIQNDVVD 109
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 59 AGVTEGITEEDQLMRAKYVYDSAFHPDTGEKVVLIGRMSAQV-----PMNMTIT 107
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 110 EGPYTCSTQTDN-----HPKT-SRVHLIVQSPKIVEISSDISINEGNISLT 156
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 108 GCMLEFYRKTPTVFWQWVNOSEFNALVNYSNRSGDTPITVRQLGTAYVSATTGAVATAG 167
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 157 -CIAT-GRDEPTVT-WRHISPKAVGFVSDEYLEIQGITREQGEYECASANDVAAPVVR 213
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 168 LKSLTGHLPPLV---GRFVPFAVAANAC 193
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 214 RVNVTVNPYPYISBAKGTGVPVGQKGTLOC 243
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
```

Search completed: February 3, 2004, 09:13:15
Job time : 19 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2004, 09:11:53 / Search time 41 Seconds
(without alignments)
1642.725 Million cell updates/sec

Title: US-09-990-415a-2

Perfect score: 1345
Sequence: 1 MESMGELPLDINIOEPRMD.....PMAIPLIMDTLEKDFLK 261

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

SPTREMBL_23: *
1: sp_archaea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phase: *
10: sp_plant: *
11: sp_rodent: *
12: sp_virus: *
13: sp_vertebrate: *
14: sp_unclassified: *
15: sp_virus: *
16: sp_bacteriap: *
17: sp_archaeap: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1345	100.0	325	4	Q8NCJ0
2	1063.5	79.1	288	11	Q8C1Z2
3	754	56.1	321	5	Q9VNI3
4	742.5	55.2	324	5	Q22161
5	693	51.5	326	5	Q18666
6	615.5	45.8	327	5	Q9VWV3
7	614.5	45.7	327	5	Q8S296
8	582.5	43.3	349	5	Q17863
9	508.5	37.8	319	5	Q09972
10	497	37.0	201	5	Q45976
11	484.5	36.0	340	4	Q8TD22
12	484.5	36.0	342	11	Q8CFD0
13	484.5	36.0	342	11	Q8BGG3
14	483.5	35.9	302	11	Q8BRQ9
15	464.5	34.5	327	3	Q12029
16	411.5	30.6	325	3	Q13793

17	320.5	23.8	107	5	Q77218	Q77218 scapitrosc
18	315	23.4	346	5	Q9NE63	Q9NE63 leishmania
19	311	23.1	144	11	Q8BPC9	Q8BPC9 mus musc
20	278	20.7	326	5	Q94603	Q94603 leishmania
21	101	7.5	791	16	Q9BRJ1	Q9BRJ1 mycoplasma
22	94.5	7.0	412	3	Q43029	Q43029 schizosacch
23	90.5	6.7	727	10	Q8RZD8	Q8RZD8 oryza sativ
24	89.5	6.7	1134	16	Q9RJU0	Q9RJU0 streptomyce
25	89	6.6	344	8	Q8SIU5	Q8SIU5 xanthusia he
26	89	6.6	650	16	Q8CQ21	Q8CQ21 streptococ
27	88.5	6.6	344	11	Q8BG33	Q8BG33 mus muscu
28	88	6.5	380	17	Q97C55	Q97C55 thermoplas
29	88	6.5	509	16	Q8G741	Q8G741 bifidobacte
30	87.5	6.5	569	16	Q9X219	Q9X219 rhizobium m
31	87.5	6.5	796	16	Q9RL24	Q9RL24 streptomyce
32	87	6.5	343	16	Q8YTX7	Q8YTX7 anabaena sp
33	87	6.5	543	16	Q8EJG7	Q8EJG7 shevanelia
34	87	6.5	641	16	Q8PEL8	Q8PEL8 xanthomonas
35	86.5	6.4	729	2	Q8RNN4	Q8RNN4 legionella
36	86	6.4	343	8	Q8WES1	Q8WES1 ceratophora
37	86	6.4	436	16	Q8XEM4	Q8XEM4 salmonella
38	85.5	6.4	301	15	Q9WHI9	Q9WHI9 walleye epi
39	85	6.3	458	17	P95947	P95947 sulfolobus
40	85	6.3	639	2	Q8VPC0	Q8VPC0 treponema d
41	84.5	6.3	574	16	Q8XQ27	Q8XQ27 talitronia s
42	84.5	6.3	756	16	Q8K5N0	Q8K5N0 streptococ
43	84.5	6.3	778	16	Q99X55	Q99X55 streptococ
44	84.5	6.3	778	16	Q8N261	Q8N261 streptococ
45	84	6.2	291	16	Q9Z4Y9	Q9Z4Y9 streptomyce

ALIGNMENTS

RESULT 1

ID Q8NCJ0 PRELIMINARY; PRT; 325 AA.

DT 01-OCT-2002 (TREMBLrel. 22, Created)

DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)

DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)

DE Hypothetical protein FLJ90226.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

NCBI_Taxid=9606;

RP SEQUENCE FROM N.A. Isegai T., Oca T., Nishikawa T., Hayaishi K., Otsuki T., Sugiyama T.,

RA Isegai T., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,

RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahari K.,

RA Masuho Y., Ono T., Okano K., Yoshikawa Y., Aotaka S., Sasaki N.,

RA Hattori A., Okumura K., Iwayanagi T., Minomiya K.,

RT "NEDO human cDNA sequencing project";

RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AK074707; BAC11151.1; ..

DR InterPro; IPR004686; MCC.

DR Pfam; PF03820; MCC; 1.

DR TIGRPFAM; TIGR00798; MCC; 1.

SW Hypothetical protein.

SEQUENCE 325 AA; 35979 MW; 783F288DAF52C137 CRC64;

Query Match 100.0%; Score 1345; DB 4; Length 325;

Best Local Similarity 100.0%; Pred. No. 1.1e-114;

Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESMGELPLDINIOEPRMDSTFLGRARHFFVTYTPRNLISGAQLASRNITQNYAG 60

DB 1 MESMGELPLDINIOEPRMDSTFLGRARHFFVTYTPRNLISGAQLASRNITQNYAG 60

QY 61 VVTGITEEDLWRKAYVDSAFHPTDGEKVLIRMSAQVPMNNTTSCMLTFYRKPFTV 120

DB 61 VVTGITEEDLWRKAYVDSAFHPTDGEKVLIRMSAQVPMNNTTSCMLTFYRKPFTV 120

```

QY 121 VFWQWVNSFNAIVNSRSGDPTITVROLGTAIVASATTGAVATAGLKSLLTKHLPPLVG 180
DB 121 VFWQWVNSFNAIVNSRSGDPTITVROLGTAIVASATTGAVATAGLKSLLTKHLPPLVG 180
QY 181 RFPVPAVAALANCINIPLMKORELQVGPVADENAGRLGYSVTAAKGIPOVVISRICMA 240
DB 181 RFPVPAVAALANCINIPLMKORELQVGPVADENAGRLGYSVTAAKGIPOVVISRICMA 240
QY 241 IPMAIPLIMDTLEKKDFLK 261
DB 241 IPMAIPLIMDTLEKKDFLK 261

RESULT 2
08C122 PRELIMINARY; PRT; 288 AA.
AC 08C122;
DT 01-MAR-2003 (TEMBLrel. 23, Created)
DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE Sideroflexin 3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium;
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
SQ EMBL: AK089885; BAC41029.1;
SQ SEQUENCE 288 AA; 31630 MW; E385501849DECC2A CRC64;

Query Match 79.14; Score 1063.5; DB 11; Length 288;
Best Local Similarity 82.9%; Pred. No. 5,1e-89;
Matches 213; Conservative 6; Mismatches 5; Indels 33; Gaps 1;

QY 5 MGBLPDINIOERWDOSTLGRARRHPTTDPNNLLSGAQLFASNTYQNTAGVYTP 64
DB 1 MGBLPDINIOERWDOSTLGRARRHPTTDPNNLLSGAQLFASNTYQNTAGVYTP 60
QY 65 GITEDOLMARAKYVDSAFHPDTGKVVLLIGMSAQVPMNTTTCMLTFPKPTVFWQ 124
DB 61 GITEDOLMARAKYVDSAFHPDTGKVVLLIGMSAQVPMNTTTCMLTFPKPTVFWQ 110
QY 125 WVAQSFNAIVNSRSGDPTITVROLGTAIVASATTGAVATAGLKSLLTKHLPPLVG 184
DB 111 -----RQLGTAIVASATTGAVATAGLKSLLTKHLPPLVG 147
QY 185 FAVAALANCINIPLMKORELQVGPVADENAGRLGYSVTAAKGIPOVVISRICMA 244
DB 148 FAVAALANCINIPLMKORELQVGPVADENAGRLGYSVTAAKGIPOVVISRICMA 207
QY 245 AIPPLIMDTLEKKDFLK 261
DB 208 AIPPLIMDTLEKKDFLK 224

RESULT 3
09VNI3 PRELIMINARY; PRT; 321 AA.
AC 09VNI3;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE CG11739 protein (AT24389p) (RH48017p).
GN CG11739.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

```

```

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananthides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Branton R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballwey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostali M., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
RA Jallat M., Kalush F., Kapen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laeko P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshire A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Rainert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson W., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Sytkas R., Teclor C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Weissman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Aghayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guatin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacle J., Parasas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Y;
RA Stapleton M., Brokstein P., Hong L., Aghayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guatin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacle J., Parasas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
AC 09VNI3;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE CG11739 protein (AT24389p) (RH48017p).
GN CG11739.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

```

[illegible]

	RESULT 4	022161	
ID.	Q22161	PRELIMINARY;	PRT; 324 AA.
AC	Q22161;		
DT	01-NOV-1996	(TREMBLrel. 01, Created)	
DT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)	
DT	01-MAR-2003	(TREMBLrel. 23, Last annotation update)	
DE	T04F8.1 protein.		
CN	T04F8.1.		
OS	Caenorhabditis elegans.		
OC	Eukaryota; Metazoa; Nematoda; Chromadorea;		
NC	Rhabdittidae; Pelododerinae; Caenorchabditae.		
OX	[1] Taxid=6239;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Lemard N.;		
RL	Submitted (NOV-1995) to the EMBL/genbank/DBJ databases.		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=99069613; Pubmed=9851916;		
RA	none;		
RT	"Genome sequence of the nematode C.elegans: A platform for		
RT	investigating biology."		
RL	Science 282:2012-2018(1998).		
DR	EMBL; Z65655; CA91477.1; -.		
DR	MormPep; T04F8.1; CE03614.		
DR	InterPro: IPR004686; Mcc.		
DR	Pfam; PF03820; Mcc; 1.		
DR	TIGRFAMs; TIGR00798; mtc; 1.		
SQ	SEQUENCE 324 AA; 35488 MW; 8203BD51E818C545 CRC64;		
	Query Match	55.2%;	Score 742.5; DB 5; Length 324;
	Best Local Similarity	55.6%;	Pred No. 14e-59;
	Matches 145; Conservative 43; Mismatches 72; Indels 1; Gaps 1		
Dy	1 MESKMGELPLDINIOEPWDOSTFLGARAHFTVTDPRTLISGAOLEASRNIVONTYAG 60		
Dz	1 MSELVKTLLVPIDIEKPKWDGTFEGRAKHFPALINPLLFGEKQLDEPFKIIVEDYRK 60		
Dy	61 VTPESITDDOLMRACKYYVDSAFHPDTGEKVLIIGMSAOVPMNNMTTTCMLTFYKTIPTV 120		
Dz	61 SVSNDLTLLNQMKAKGVVDSAFHPBTGEKMMNVGRMSAQPVMNMATGGMLTFYKS PMAY 120		
Dy	121 VFMOVNVSFNAINVYSNRSGDTPITVRQLTAAYVASATGAVATAIGLSIKHLPLVYG 180		
Dz	121 IFMQVLNSFNANVVYTNRSDDGGS-VSGQLLVSYCAATIGALLTAIGLNLVKKAPPLVG 179		
Dy	181 RFVPEAAVAANANCINIPLMRORELGVGI PVADDAQRLGYSTAARKGIFOVVISRICMA 240		

Db	180	KLVPFAVCVANSINIPAMRGELEGGIDIDDENQVITIGSPVAQSAISQVVRIFMA	239
Qy	241	IPMAIPPLIMDLKKOFLK	261
Db	240	VPSFAFIPVVNALEKRPYFK	260

ID	Q18666	PRELIMINARY;	PRT;	326 AA.
AC	Q18666			
DT	01-NOV-1996	(TREMBLrel. 01, Created)		
DT	01-JAN-1998	(TREMBLrel. 05, Last sequence update)		
DT	01-MAR-2003	(TREMBLrel. 23, Last annotation update)		
DE	C47D12.3 protein.			
GN	C47D12.3.			
OS	Caenorhabditis elegans.			
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;			
OC	Rhabditidae; Peloderinae; Caenorhabditis.			
OX	NCBI_TaxId=6239;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Gajadacy S.,			
RL	Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=99069613; PubMed=9851916;			
RA	none;			
RT	"Genome sequence of the nematode C.elegans: A platform for			
RT	investigating biology."			
RL	Science 282:2012-2018(1998).			
DR	EMBL; Z69902; CA93764.1; -.			
DR	Wormped; C47D12.3; CB05431.			
DR	InterPro; IPR004686; MTC.			
DR	Pfam; PF03820; MTC; 1.			
DR	TIGRFAMs; TIGR00798; mtc; 1.			
DR	SEQUENCE 326 AA; 36436 MW; 8D0F2C3B604F7C27 CRC64;			
SO	SEQUENCE			
QY	3 SKMGELPLDINIOBPRWDOSTPLGRARHFTVTDPRMLLSGAOLEASRNIVONTAGV 62			
Db	6 SKCTELP---DISPRWDQNTFGGRVNVFFETACNCLFVSNALKLEKARNIVLEVKQKCY 62			
QY	63 TPGITTEQLRAKAVYDSAFHEDPTEGEKVLIGRNSAOVPMNMNTITGCMLTTRYKTPYVF 122			
Db	63 DPNMTVELMWAKTLVYSAPFHPDTEGEKWLIFGRMSAOVPCMLLTGCMLTTRYQKLPHVIF 122			
QY	123 WQWNOSFNMIWVNSNRSGDPTPIVROLGSTAYVSATTGAVATALGKSLT---KHLPLPV 179			
Db	123 FHWNOSFNALVWNTVNSNGTHKKODRLLISYCGHTTGLSCALSPNMLKKKNAPIL 182			
QY	180 GRFPVFAVAANACINIPLMKORELQVIGIPVADAGORLGSVTAAMKOGIFQVVISRICM 239			
Db	183 ARLVPPAIAANANINIPMNKPKFTNGIPIPEDGGRIMGRSVTAAPHALIQVVLRSVGM 242			
QY	240 AIPMAIIPRLIMDTLEK 256			
Db	243 AVPMNVLGPVILBOLSK 259			
RESULT 6				
ID	Q9VWV3	PRELIMINARY;	PRT;	327 AA.
AC	Q9VWV3			
DT	01-MAY-2000	(TREMBLrel. 13, Created)		
DT	01-OCT-2002	(TREMBLrel. 22, Last sequence update)		
DT	01-MAR-2003	(TREMBLrel. 23, Last annotation update)		
DE	CG6812 protein.			
GN	CG6812.			

SQ SEQUENCE 327 AA; 36468 MW; AED2357693D50B92 CRC64;
 Query Match 45.7%; Score 614.5; DB 5; Length 327;
 Best Local Similarity 50.6%; Pred. No. 7e-48;
 Matches 128; Conservative 45; Mismatches 79; Indels 1; Gaps 1;
 QY 5 MGEPLDINIDPRDOSTFLGRANHFPTVDPRLNLISGAOLEASRIIVQNYRAGVTP 64
 Db 1 MSQVSTLIDVDKPLFDLSTFAGRFQYFAMMTDPRIVVSSDRLLKAKAMVRKYRGDOSP 60
 QY 65 GITEDELRAKYVDSAFHPTGEEKVNLIGRMSAQPVMNMTITGCMLEFYRKTPTVVFMO 124
 Db 61 PLKPEBYVMNKLINSAFHPDTGELQNFGRNSFQVPGMLITGMLAFYRTVPVAVLMQ 120
 QY 125 WVNOSFNAIVNYSNRSGDPTTVRQLGTAVYSATTGAVALGLKSL-TKHLPLVGRFV 183
 Db 121 FINOSFNAVVMYTRNNANSPSTVQLGVAVYSATTGALVAALIGCNVYSKATPLFQGFV 180
 QY 184 PFAVVAANNCINIPLMRORELOVGI PVADGAGRLGYSVTAAKQIFOVVISRICMAIPA 243
 Db 181 PFAVVAANFNIPILMRQNEIINGIEVKSDDGVVVGOSRLAIKIGEVVSRIMAAFG 240
 QY 244 MAIPPLIMDTLEK 256
 Db 241 MLVDEPLIMERLEK 253

RESULT 8

ID 017863 PRELIMINARY; PRT; 349 AA.
 AC 017863
 DT 01-JAN-1998 (TRENBLrel. 05, Created)
 DT 01-OCT-2001 (TRENBLrel. 18, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE F37H8.4 protein.
 GN F37H8.4
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RN SEQUENCE FROM N.A.
 RA Gregory J.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 RP [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C. elegans: A platform for
 .rt investigating biology."
 RL Science 282:2012-2018(1998).
 DR EMBL; Z81534; CAB04347.2; -;
 DR WormPep; F37H8.4; CE28302.
 DR InterPro; IPR004686; MCC.
 DR Pfam; PF03820; MCC; 1.
 DR TIGRFAMs; TIGR00798; MCC; 1.
 SQ SEQUENCE 349 AA; 39351 MW; 72CA25828404EF82 CRC64;
 Query Match 43.3%; Score 582.5; DB 5; Length 349;
 Best Local Similarity 43.4%; Pred. No. 6.5e-45;
 Matches 121; Conservative 47; Mismatches 86; Indels 25; Gaps 2;
 QY 1 MESKMGELPLDINIOEPRDOSTFLG-----RAHFEFTVDP 37
 Db 1 MSEIVSNLKRTPDISKPMQSQRTYVGNKSSGQKVPAKRLFRMDVSESRVRHFFTLINP 60
 QY 38 RNLLLSGAOLEASRIIVQNYRAGVTPGITEQLMRKAVYVDSAFHPRTGKVVILIGMS 97
 Db 61 LTLTSSVARQOQCOIYLDYNGKGSPTLVSELMKATTLVDSYHPDTGKMFELGMS 120
 QY 98 AQVPMNMTITGCMLEFYRKTPTVVFMOVWNOSFNAIVNYSNRSGDPTTVRQLGTAVYSA 157
 Db 121 AQVPMNMTITGCMLEFYRKTPTVVFMOVWNOSFNAIVNYSNRSGDPTTVRQLGTAVYSA 180

QY 158 TTGAVALGLKSLTKHLPLVGRFVPPAAVAANNCINIPLMRORELOVGI PVADGAGOR 217
 Db 181 TGGAMAAALSLNANKN--SIAELVFPFAVALANTINIPMIRSNVEYEGELRDENGEL 238
 QY 218 LGYSTAAKQIFOVVISRICMAIPMAIPLIMDTLEK 256
 Db 239 LARSQMAIISIAQVTLISRIAMAMPDMVMTPIINMRITR 277

RESULT 9

ID 009972 PRELIMINARY; PRT; 319 AA.
 AC 009972
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Hypothetical 35.3 kDa protein.
 GN C14F5.4.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=Bristol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA Waterston R.;
 RT "Genome sequence of the nematode C. elegans: a platform for
 .rt investigating biology. The C. elegans Sequencing Consortium."
 RL Science 282:2012-2018(1998).
 RN [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=Bristol N2;
 RA Minx P.;
 RT "The sequence of C. elegans cosmid C14F5."
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RN SEQUENCE FROM N.A.
 RP STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U29082; AAA68404.3; -;
 DR WormPep; C14F5.4; CE30370.
 DR InterPro; IPR004686; MCC.
 DR Pfam; PF03820; MCC; 1.
 DR TIGRFAMs; TIGR00798; MCC; 1.
 KM Hypothetical protein.
 SQ SEQUENCE 319 AA; 35323 MW; B400FEA731B27F93 CRC64;
 Query Match 37.8%; Score 508.5; DB 5; Length 319;
 Best Local Similarity 44.4%; Pred. No. 3.3e-38;
 Matches 108; Conservative 39; Mismatches 93; Indels 3; Gaps 2;
 QY 16 EPRWDOSTFLGRANHFPTVDPRLNLISGAOLEASRIIVQNYRAGVTPGITEQLMRK 75
 Db 7 QPRYDOSTFLGRANHFPTVDPRLNLISGAOLEASRIIVQNYRAGVTPGITEQLMRK 65
 QY 76 YVDSAFHPTGEEKVNLIGRMSAQPVMNMTITGCMLEFYRKTPTVVFMOVWNOSFNAIVN 135
 Db 66 RLCSAHPDTGELQNFGRNSFQVPGMLITGMLAFYRTVPVAVLMQ 125
 QY 136 YSNRSGDPTTVRQLGTAVYSATTGAVALGLKSL-LTKHLPLVGRFVPPAAVAANNC 193
 Db 126 YTRNNASTLTITDLDVVSYSYTAVALAMALGLKTYPAKKQSSPLAQLVPLGAVALANA 185
 QY 194 INIPLMRORELOVGI PVADGAGRLGYSVTAAKQIFOVVISRICMAIPMAIPLIMDT 253
 Db 186 INIPMKNQELKQGMVTLADGNVGVSRLLAAKALVVLISRIIVAPCMILTPVMEG 245
 QY 254 LEK 256
 Db 246 LNK 248

ID	Q045976	PRELIMINARY;	PRT;	201 AA.
AC	Q045976	PRELIMINARY;	PRT;	201 AA.
DT	01-JUN-1998	(TREMBLrel. 06, Created)		
DT	01-MAR-2003	(TREMBLrel. 23, Last sequence update)		
DT	01-MAR-2003	(TREMBLrel. 23, Last annotation update)		
DE	Hypothetical protein Y6E2A.9.			
GN	Y6E2A.9			
OS	Caenorhabditis elegans.			
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;			
OC	Rhabditidae; Peleoidae; Caenorhabditis.			
OX	NCBI_TaxID=6239;			
RN	[1]			
RA	SEQUENCE FROM N.A.			
RL	Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.			
RP	[2]			
RX	SEQUENCE FROM N.A.			
RX	MEDLINE=99069613; PubMed=9851916;			
RA	none;			
RT	"Genome sequence of the nematode C.elegans: A platform for			
RT	investigating biology."			
RL	Science 282:2012-2018 (1998).			
DR	EMBL; AL021175; CAA15970.2; -.			
KW	Hypothetical protein.			
SO	SEQUENCE 201 AA; 22446 MW; BFB5B1CEBF4E6C4C CRC64;			
Query Match	37.0%; Score 497; DB 5; Length 201;			
Best Local Similarity	50.0%; Pred. No. 2e-37;			
Matches	93; Conservative 29; Mismatches 64; Indels 0; Gaps 0;			
QY	13 NIOEPRWQSTFLGRARHFFVTDPRLNLLSGAOLEASRNIVQYRAGVTPGITEPOLM 72			
DB	13 DISKSKLDLDTYSRGVRYFPASANMTLPTSSNTGEMCRKIVDYKKGINPELIMDELN 72			
QY	73 RAKTYVDSAFHPDTEGKRVLLIGRMSAQPVMNMTITGCMLTFYRKTPTVPEWQVNSFNA 132			
DB	73 SAKLIYDSVHPDTEGKRFKFCGRMSAQPPANVITGMLISCYRTPGILTFSHINQSFNA 132			
QY	133 IVVNSNRSGDPIPTVROLGTAIVYSATITGAVALALKSLTGLPPLNQRFPVPAVAAN 192			
DB	133 IVVNTISNGNCRTTNOQLLYSTFCATGAATTAALGLNMMVKNSHGLAGLVPEVAAVAN 192			
QY	193 CINIPL 198			
DB	193 AINIPM 198			
RESULT 11				
ID	Q08TD22	PRELIMINARY;	PRT;	340 AA.
AC	Q08TD22;			
DT	01-JUN-2002	(TREMBLrel. 21, Created)		
DT	01-JUN-2002	(TREMBLrel. 21, Last sequence update)		
DT	01-OCT-2002	(TREMBLrel. 22, Last annotation update)		
DE	Sideroflexin 5.			
GN	SFXN5.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RA	SEQUENCE FROM N.A.			
RA	MEDLINE=22035372; PubMed=12039050;			
RA	Lockhart P.J., Holton B., Lincoln S., Husey J., Zimprich A.,			
RA	Gasser T., Wszolek Z.K., Hardy J., Farrer M.J.;			
RT	"The human sideroflexin 5 (SFXN5) gene: sequence, expression analysis			
RT	and exclusion as a candidate for PARK3."			
RL	Gene 285:229-237 (2002).			
OR	EMBL; AY044437; AAK95826.1; -.			

[illegible]

Db 211 ASANICNVLMRYGELGIVLDADGVLVSSKIAARHALLFTRVLVPMFIVLP 270
 Qy 249 LIMDTLEKDFLK 261
 Db 271 IVMSMLEKTALLQ 283

RESULT 13

Q8BG3 PRELIMINARY; PRT; 342 AA.
 ID Q8BG3;
 AC Q8BG3;
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Sideroflexin 5.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium.
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs."
 RT Nature 420:563-573 (2002).
 RL EMBL; AK036285; BAC29372.1; -
 DR EMBL; AK082073; BAC38403.1; -
 SQ SEQUENCE 342 AA; 37328 MW; 9BDB8261D5EF79D9 CRC64;

Query Match 36.0%; Score 484.5; DB 11; Length 342;
 Best Local Similarity 38.3%; Pred. No. 5.8e-36;
 Matches 97; Conservative 50; Mismatches 99; Indels 7; Gaps 2;

Qy 16 EPRNDQSTFLGRARHFTVDPRLILSGAQLSRIIVYRAGVTPGTTEQLRPAK 75
 Db 31 KPRQQTSPFYGRFHFDDIDPRTLFTYKRLRAVOLLEBYKKGTLRPGVTNQLSAQ 90
 Qy 76 VYVDSAPHPTGKRVNLIGRMSAQVPMNMITGCMLEFYRKTPTVPMQVNOSEFNAIVN 135
 Db 91 KIKQAIHPDTNEKIFPMFPMISGTFPGTPIVGLLPMQTLASTVPMQINOSHACVN 150
 Qy 136 YSNRSGDPTITVRQLGTAIVSATTGAVATLGLKSL---TKHLPP--LVGRFVPPAAV 188
 Db 151 YANRNATKPSPAKFIQGYLGAVISAVGLNVLVQKANKFPATRLVQGRFVPPAV 210
 Qy 189 AAANCINIPLMKORELOVGIIVADAGORLGYSTAAKQIFQVVISRICAIIPMAIIP 248
 Db 211 ASANICNVLMRYGELGIVLDADGVLVSSKIAARHALLFTRVLVPMFIVLP 270
 Qy 249 LIMDTLEKDFLK 261
 Db 271 IVMSMLEKTALLQ 283

RESULT 14

Q8BR09 PRELIMINARY; PRT; 302 AA.
 ID Q8BR09;
 AC Q8BR09;
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Sideroflexin 5.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Cortex;

RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium.
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs."
 RT Nature 420:563-573 (2002).
 RL EMBL; AK043706; BAC31624.1; -
 DR EMBL; AK082073; BAC38403.1; -
 SQ SEQUENCE 302 AA; 32783 MW; 650F6A97043EF26 CRC64;

Query Match 35.9%; Score 483.5; DB 11; Length 302;
 Best Local Similarity 38.6%; Pred. No. 6e-36;
 Matches 96; Conservative 50; Mismatches 96; Indels 7; Gaps 2;

Qy 16 EPRNDQSTFLGRARHFTVDPRLILSGAQLSRIIVYRAGVTPGTTEQLRPAK 75
 Db 31 KPRQQTSPFYGRFHFDDIDPRTLFTYKRLRAVOLLEBYKKGTLRPGVTNQLSAQ 90
 Qy 76 VYVDSAPHPTGKRVNLIGRMSAQVPMNMITGCMLEFYRKTPTVPMQVNOSEFNAIVN 135
 Db 91 KIKQAIHPDTNEKIFPMFPMISGTFPGTPIVGLLPMQTLASTVPMQINOSHACVN 150
 Qy 136 YSNRSGDPTITVRQLGTAIVSATTGAVATLGLKSL---TKHLPP--LVGRFVPPAAV 188
 Db 151 YANRNATKPSPAKFIQGYLGAVISAVGLNVLVQKANKFPATRLVQGRFVPPAV 210
 Qy 189 AAANCINIPLMKORELOVGIIVADAGORLGYSTAAKQIFQVVISRICAIIPMAIIP 248
 Db 211 ASANICNVLMRYGELGIVLDADGVLVSSKIAARHALLFTRVLVPMFIVLP 270
 Qy 249 LIMDTLEK 257
 Db 271 IVMSMLEKQ 279

RESULT 15

Q12029 PRELIMINARY; PRT; 327 AA.
 ID Q12029;
 AC Q12029;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
 DE Chromosome XV reading frame ORF YOR271C.
 GN YOR271C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 NX NCBI_TaxId=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Janniaux J.C., Polley R.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Cheret G., Sor F.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA MIPs;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RA Cheret G., Bernardi A., Sor F.J.;
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; 275179; CA99495.1; -
 DR EMBL; X89633; CA61777.1; -
 DR SGD; S0005797; YOR271C.
 DR InterPro; IPR004686; Mtc.
 DR Pfam; PF03820; Mtc; 1.
 DR TIGRFAMs; TIGR00798; mtc; 1.
 KV Hypothetical protein.
 SQ SEQUENCE 327 AA; 35414 MW; FD175626E63B8619 CRC64;

Qy 241 IPMAIPPLIMDTLEKDFLK 261
Db 241 IPMAIPPLIMDTLEKDFLK 261

RESULT 2
US-09-867-550-1344
Sequence 1344, Application US/09867550
Patent No. US20020082206A1
GENERAL INFORMATION:
APPLICANT: Leach, Martin D.
APPLICANT: Mehrtad, Foad,
APPLICANT: Conley, Pamela
APPLICANT: Law, Debbie
APPLICANT: Topper, James
TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and
FILE REFERENCE: 21402-013 (Cura-113)
CURRENT FILING DATE: 2001-09-20
PRIOR FILING DATE: 2001-09-20
NUMBER OF SEQ ID NOS: 2125
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1344
LENGTH: 266
TYPE: PRT
ORGANISM: Homo sapiens
US-09-867-550-1344

Query Match 100.0%; Score 1345; DB 9; Length 266;
Best Local Similarity 100.0%; Pred. No. 2, 2e-141;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MESMGGLPLDINIOEPRWDSTFLGRARHFTVTPRLILSGOLEASRNIVONRAG 60
Db 1 MESMGGLPLDINIOEPRWDSTFLGRARHFTVTPRLILSGOLEASRNIVONRAG 60
Qy 61 VTPGTEPOLWRAKYVDSAPHPDTGEKVLIGMSAQVPMNMTTTCMLTFRKPTPV 120
Db 61 VTPGTEPOLWRAKYVDSAPHPDTGEKVLIGMSAQVPMNMTTTCMLTFRKPTPV 120
Qy 121 VFMWVNSFNALVYNSRSGDPTITVROLGTAAYVATTGAVATLGLSKLTGLPLV 180
Db 121 VFMWVNSFNALVYNSRSGDPTITVROLGTAAYVATTGAVATLGLSKLTGLPLV 180
Qy 181 RVPFAAVALANCINIPMRQRELOVGPVADGAGRLGYSVTAAKQIFOVVISRICMA 240
Db 181 RVPFAAVALANCINIPMRQRELOVGPVADGAGRLGYSVTAAKQIFOVVISRICMA 240
Qy 241 IPMAIPPLIMDTLEKDFLK 261
Db 241 IPMAIPPLIMDTLEKDFLK 261

RESULT 3
US-09-990-415a-8
Sequence 8, Application US/09990415A
Patent No. US20020165182A1
GENERAL INFORMATION:
APPLICANT: Pharmacia AB
TITLE OF INVENTION: Protein Cluster I
FILE REFERENCE: 00349
CURRENT FILING DATE: 2001-11-21
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn version 3.0
SEQ ID NO 8
LENGTH: 322
TYPE: PRT
ORGANISM: human
FEATURE:
NAME/KEY: misc_feature

LOCATION: (25)-(25)
OTHER INFORMATION: Xaa=A,T,G or C
US-09-990-415a-8

Query Match 81.0%; Score 1089; DB 10; Length 322;
Best Local Similarity 78.5%; Pred. No. 1, 1e-112;
Matches 201; Conservative 28; Mismatches 27; Indels 0; Gaps 0;
Qy 6 GELPLDINIOEPRWDSTFLGRARHFTVTPRLILSGOLEASRNIVONRAGVTTG 65
Db 3 GELPLDINIOEPRWDSTFLGRARHFTVTPRLILSGOLEASRNIVONRAGVTTG 65
Qy 66 ITEDOLWRAKYVDSAPHPDTGEKVLIGMSAQVPMNMTTTCMLTFRKPTPVFMQ 125
Db 66 ITEDOLWRAKYVDSAPHPDTGEKVLIGMSAQVPMNMTTTCMLTFRKPTPVFMQ 125
Qy 126 VNSFNALVYNSRSGDPTITVROLGTAAYVATTGAVATLGLSKLTGLPLVGRFVP 185
Db 126 VNSFNALVYNSRSGDPTITVROLGTAAYVATTGAVATLGLSKLTGLPLVGRFVP 185
Qy 186 AAVAANCINIPMRQRELOVGPVADGAGRLGYSVTAAKQIFOVVISRICMAIPMA 245
Db 186 AAVAANCINIPMRQRELOVGPVADGAGRLGYSVTAAKQIFOVVISRICMAIPMA 245
Qy 246 IPPPLIMDTLEKDFLK 261
Db 246 IPPPLIMDTLEKDFLK 261

RESULT 4
US-10-014-338-2
Sequence 2, Application US/10014338
Publication No. US2003009261A1
GENERAL INFORMATION:
APPLICANT: Herath, et al.
TITLE OF INVENTION: ADP1-41, A NOVEL PROTEIN ISOLATED FROM BRAIN TISSUE HOMOGENATE A
FILE REFERENCE: 9195-077
CURRENT FILING DATE: 2002-05-01
PRIOR FILING DATE: 2001-12-10
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 322
TYPE: PRT
ORGANISM: Homo sapiens
US-10-014-338-2

Query Match 81.0%; Score 1089; DB 15; Length 322;
Best Local Similarity 78.5%; Pred. No. 1, 1e-112;
Matches 201; Conservative 28; Mismatches 27; Indels 0; Gaps 0;
Qy 6 GELPLDINIOEPRWDSTFLGRARHFTVTPRLILSGOLEASRNIVONRAGVTTG 65
Db 3 GELPLDINIOEPRWDSTFLGRARHFTVTPRLILSGOLEASRNIVONRAGVTTG 65
Qy 66 ITEDOLWRAKYVDSAPHPDTGEKVLIGMSAQVPMNMTTTCMLTFRKPTPVFMQ 125
Db 66 ITEDOLWRAKYVDSAPHPDTGEKVLIGMSAQVPMNMTTTCMLTFRKPTPVFMQ 125
Qy 126 VNSFNALVYNSRSGDPTITVROLGTAAYVATTGAVATLGLSKLTGLPLVGRFVP 185
Db 126 VNSFNALVYNSRSGDPTITVROLGTAAYVATTGAVATLGLSKLTGLPLVGRFVP 185
Qy 186 AAVAANCINIPMRQRELOVGPVADGAGRLGYSVTAAKQIFOVVISRICMAIPMA 245
Db 186 AAVAANCINIPMRQRELOVGPVADGAGRLGYSVTAAKQIFOVVISRICMAIPMA 245
Qy 246 IPPPLIMDTLEKDFLK 261
Db 246 IPPPLIMDTLEKDFLK 261